

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
6 January 2005 (06.01.2005)

PCT

(10) International Publication Number
WO 2005/001103 A2

(51) International Patent Classification⁷: C12N 15/86,
5/10, 7/01, C07K 14/075, A61K 39/235, 48/00, C12N
15/34

WILSON, James, M. [US/US]; 1350 N. Avignon Drive,
Gladwyne, PA 19035 (US).

(21) International Application Number:
PCT/US2004/016614

(74) Agents: KODROFF, Cathy, A. et al.; Howson and How-
son, Spring House Corporate Center, P.O. Box 457, Spring
House, PA 19477 (US).

(22) International Filing Date: 15 June 2004 (15.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/465,302 20 June 2003 (20.06.2003) US
60/566,212 28 April 2004 (28.04.2004) US
60/575,429 28 May 2004 (28.05.2004) US

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier application:
US 10/465,302 (CIP)
Filed on 20 June 2003 (20.06.2003)

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI,
SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): THE
TRUSTEES OF THE UNIVERSITY OF PENN-
SYLVANIA [US/US]; 3160 Chestnut Street, Suite 200,
Philadelphia, PA 19104 (US).

Published:

— *without international search report and to be republished
upon receipt of that report*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): ROY, Soumitra
[US/US]; 240 Pugh Road, Wayne, PA 19087 (US).

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADEN-
OVIRUSES

(57) Abstract: A method for providing an adenovirus from a serotype which does not grow efficiently in a desired cell line with the ability to grow in that cell line is described. The method involves replacing the left and right termini of the adenovirus with the corresponding termini from an adenovirus which grow efficiently in the desired cell line. At a minimum, the left terminus spans the (5') inverted terminal repeat, the left terminus spans the E4 region and the (3') inverted terminal repeat. The resulting chimeric adenovirus contains the internal regions spanning the genes encoding the penton, hexon and fiber from the serotype which does not grow efficiently in the desired cell. Also provided are vectors constructed from novel simian adenovirus sequences and proteins, host cells containing same, and uses thereof.



WO 2005/001103 A2

METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES

BACKGROUND OF THE INVENTION

5 The presence of humoral immunity (circulating antibodies) to adenovirus capsid proteins is a barrier to the use of adenovirus vectors for gene therapy. The prototype adenovirus vectors that have been developed for gene therapy are based on subgroup C adenoviruses such as that of serotype 5. The prevalence of neutralizing antibodies against subgroup C adenoviruses is generally high in human populations as
10 a result of frequent exposure to these pathogens. This fact is likely to greatly limit the effectiveness of gene therapy vectors based on serotypes such as Ad5.

 Analysis of the nature of the protective antibodies against adenoviruses has indicated that the most important target is the major capsid protein, hexon [Wolfhart (1988) *J. Virol.* 62, 2321; Gall *et al.* (1996) *J. Virol.* 70, 2116]. Several efforts have
15 been made to engineer the hexon so as to evade the anti-hexon antibodies by making chimeric adenoviruses harboring hexons from other serotypes [Roy *et al.* (1998) *J. Virol.* 72, 6875; US Patent No 5,922,315; Gall *et al.* (1998) *J. Virol.* 72, 10260; Youil *et al.* (2002) *Hum. Gene Ther.* 13, 311; Wu *et al.* (2002) *J. Virol.* 76, 12775].
 However, this has been largely unsuccessful when exchanges among distant serotypes
20 are attempted.

 Alternatively, investigators have proposed using adenovirus vectors that rarely cause human infections or using adenoviruses from non-human sources. However, the lack of a practical manner in which to produce large numbers of such vectors has proved to be a hindrance to developing such vectors.

25

SUMMARY OF THE INVENTION

 The present invention provides a method of modifying adenoviruses having capsids, and particularly, including hexons, from serotypes which are not well
30 adapted for growth in cells useful for adenoviral virion production. The method is useful for production of scalable amounts of adenoviruses. The resulting chimeric adenovirus genomes are useful for a variety of purposes which are described herein.

The invention further provides novel, isolated, adenovirus SA18 nucleic acid and amino acid sequences, vectors containing same, cell lines containing such SA18 sequences and/or vectors, and uses thereof.

Other aspects and advantages of the present invention will be readily apparent
5 from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 provides the map of the genome of the simian adenovirus generated by shotgun cloning as described in the examples below.

10 Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus, termed H5S25H5.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides chimeric adenovirus genomes composed of
15 the left terminal end and right terminal end of an adenovirus which can be cultured in the selected host cell, and the internal regions encoding, at a minimum, the capsid proteins of another adenovirus serotype. This invention is particularly advantageous for generating adenoviruses having serotypes which are difficult to culture in a desired cell type. The invention thus permits generation of chimeric adenoviruses
20 vectors of varying serotypes.

In the embodiments illustrated herein, chimeric adenoviruses have been constructed where most structural proteins, and not merely the hexon or fiber, are derived from an adenovirus of an unrelated serotype, thereby preserving the majority of the protein-protein interactions that are involved in capsid assembly. Most of the
25 early genes such as those encoded by the adenovirus E1 and E4 regions that are responsible for transcription regulation and regulation of the host cell cycle, are retained from a different serotype that is known to result in high titer virus generation in the commonly used cell types, such as HEK 293 which supplies the Ad5 E1 proteins in *trans*.

30 In another embodiment, the invention provides novel nucleic acid and amino acid sequences from Ad SA18, which was originally isolated from vervet monkey [ATCC VR-943]. The present invention further provides novel adenovirus vectors and packaging cell lines to produce those vectors for use in the *in vitro* production of

recombinant proteins or fragments or other reagents. The invention further provides compositions for use in delivering a heterologous molecule for therapeutic or vaccine purposes. Such therapeutic or vaccine compositions contain the adenoviral vectors carrying an inserted heterologous molecule. In addition, novel sequences of the
5 invention are useful in providing the essential helper functions required for production of recombinant adeno-associated viral (AAV) vectors. Thus, the invention provides helper constructs, methods and cell lines which use these sequences in such production methods.

The term "substantial homology" or "substantial similarity," when referring to
10 a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences.

The term "substantial homology" or "substantial similarity," when referring to
15 amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, or a fragment thereof which is at least 8 amino acids, or more
20 desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison
25 may be over the full-length of the genome (*e.g.*, about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, *e.g.*, the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, *e.g.*, of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to
30 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least

about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full
5 length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as
10 "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and
15 percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similarly programs are available for performing
20 amino acid alignments. Generally, these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

25 As used throughout this specification and the claims, the term "comprise" and its variants including, "comprises", "comprising", among other variants, is inclusive of other components, elements, integers, steps and the like. The term "consists of" or "consisting of" are exclusive of other components, elements, integers, steps and the like.

30 Except where otherwise specified, the term "vector" includes any genetic element known in the art which will deliver a target molecule to a cell, including, naked DNA, a plasmid, phage, transposon, cosmids, episomes, viruses, etc.

By "minigene" is meant the combination of a selected heterologous gene and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

As used herein, the term "transcomplement" refers to when a gene (gene product) of one adenovirus serotype supplies an adenovirus serotype lacking this gene (gene product) from another serotype with the missing function. For example, human adenovirus serotype 5 E1a and E1b functions are known to transcomplement E1-deleted chimpanzee adenovirus Pan 9. Similarly, the inventors have found that human Ad5 E1 transcomplements E1-deleted chimpanzee adenovirus serotypes Pan5, Pan6, Pan7, and simian adenovirus serotypes SV1, SV25 and SV39. Other examples of transcomplementing serotypes include human Ad5 and human Ad2, Ad3, Ad4, Ad5, Ad7, and Ad12.

The term "functionally deleted" or "functional deletion" means that a sufficient amount of the gene region is removed or otherwise damaged, *e.g.*, by mutation or modification, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed. Other suitable sites for gene disruption or deletion are discussed elsewhere in the application.

The term "functional" refers to a product (*e.g.*, a protein or peptide) which performs its native function, although not necessarily at the same level as the native product. The term "functional" may also refer to a gene which encodes and from which a desired product can be expressed.

I. Chimeric Adenoviral Vectors

The compositions of this invention include chimeric adenoviral vectors that deliver a heterologous molecule to cells. For delivery of such a heterologous molecule, the vector can be a plasmid or, preferably, a chimeric adenovirus. The chimeric adenoviruses of the invention include adenovirus DNA from at least two source serotypes, a "donating serotype" and a "parental adenovirus" as described in more detail herein, and a minigene.

Because the adenoviral genome contains open reading frames on both strands, in many instances reference is made herein to 5' and 3' ends of the various regions to avoid confusion between specific open reading frames and gene regions. Thus, when

reference is made herein to the “left” and “right” end of the adenoviral genome, this reference is to the ends of the approximately 36 kb adenoviral genome when depicted in schematic form as is conventional in the art [see, *e.g.*, Horwitz, “Adenoviridae and Their Replication”, in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)]. Thus, as used

5 herein, the “left terminal end” of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme left end of the schematic. Typically, the left end refers to be portion of the genome beginning at map unit 0 and extending to the right to include at least the 5’ inverted terminal repeats (ITRs), and excludes the internal regions of the

10 genome encoding the structural genes. As used herein, the “right terminal end” of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme right end of the schematic. Typically, the right end of the adenoviral genome refers to be portion of the genome ending at map unit 36 and extending to the left to include at least the 3’

15 ITRs, and excludes the internal regions of the genome encoding the structural genes.

A. Adenovirus Regulatory Sequences

1. Serotype

The selection of the adenovirus serotype donating its left terminal end and right terminal end can be readily made by one of skill in the art from

20 among serotypes which can readily be cultured in the desired cell line. Among other factors which may be considered in selecting the serotype of the donating serotype is compatibility with the adenovirus serotype which will be supplying the internal regions at the location at which their sequences are hybridized.

Suitable adenoviruses for donating their left and right termini

25 are available from the American Type Culture Collection, Manassas, Virginia, US (ATCC), a variety of academic and commercial sources, or the desired regions of the donating adenoviruses may be synthesized using known techniques with reference to sequences published in the literature or available from databases (*e.g.*, GenBank, etc.). Examples of suitable donating adenoviruses include, without limitation, human

30 adenovirus serotypes 2, 3, 4, 5, 7, and 12, and further including any of the presently identified human types [see, *e.g.*, Horwitz, “Adenoviridae and Their Replication”, in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)] which can be cultured in the desired cell. Similarly adenoviruses known to infect non-human primates (*e.g.*, chimpanzees,

rhesus, macaque, and other simian species) or other non-human mammals and which grow in the desired cell can be employed in the vector constructs of this invention. Such serotypes include, without limitation, chimpanzee adenoviruses Pan 5 [VR-591], Pan6 [VR-592], Pan7 [VR-593], and C68 (Pan9), described in US Patent No. 5 6,083,716; and simian adenoviruses including, without limitation SV1 [VR-195]; SV25 [SV-201]; SV35; SV15; SV-34; SV-36; SV-37, and baboon adenovirus [VR-275], among others. The sequences of Pan 5 (also termed C5), Pan 6 (also termed C6), Pan 7 (also termed C7), SV1, SV25, and SV39 have been described [WO 03/046124, published 5 June 2003; and in US Patent Application No. 10/739,096, 10 filed December 19, 2003)], which are incorporated by reference. In the following examples, the human 293 cells and adenovirus type 5 (Ad5), Pan9, and Ad40 are used for convenience. However, one of skill in the art will understand that other cell lines and/or comparable regions derived from other adenoviral strains may be readily selected and used in the present invention in the place of (or in combination with) 15 these serotypes.

2. Sequences

The minimum sequences which must be supplied by the adenovirus donating its left terminal end and its right terminal end include the 5' cis-elements and the 3' cis-elements necessary for replication and packaging. Typically, 20 the 5' cis-elements necessary for packaging and replication include the 5' inverted terminal repeat (ITR) sequences (which functions as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The right end of the adenoviral genome includes the 3' cis-elements (including the ITRs) 25 necessary for packaging and encapsidation. Desirably, the adenovirus serotype donating its left and right termini and/or an adenovirus serotype which transcomplements the serotype of the donating adenovirus, further provides the functions of the necessary adenovirus early genes, including E1 (E1a and E1b), E2 (E2a and E2b), and E4 (including at least the ORF6 region). E3 is not essential and 30 may be deleted as desired, *e.g.*, for insertion of a transgene in this region or to provide space for a transgene inserted in another region (typically for packaging it is desirable for the total adenoviral genome to be under 36 kb).

In certain embodiments, the necessary adenovirus early genes are contained in the chimeric construct of the invention. In other embodiment, one or more of the necessary adenovirus early genes can be provided by the packaging host cell or in *trans*.

5 In general, the chimeric adenovirus of the invention contains regulatory sequences from the donating adenovirus serotype, or a transcomplementing serotype, to provide the chimeric adenovirus with compatible regulatory proteins. Optionally, one or more of the necessary adenoviral structural genes is provided by the adenovirus donating its left terminal and its right terminal end.

10 In certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the adenovirus serotype donating its left and right termini. Where all of these genes are
15 derived from the adenovirus serotype donating the 5' and 3' ITRs, a "pseudotyped" virus is formed. In one embodiment, the chimeric adenovirus contains the left end of the adenovirus genome from the donating serotype, from the 5' ITR through the end of the pol gene (or the pTP). In another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, from the 5' ITR through the
20 penton. In yet another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, *e.g.*, through the end of pTP, but contains an ITR from an adenovirus serotype heterologous to the donating adenovirus serotype. Still other embodiments will be readily apparent from the present disclosure.

Optionally, one or more of the genes can be hybrids formed
25 from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins (*e.g.*, without limitation, polymerase, terminal protein, IIIa protein). Suitably, these genes express functional proteins which permit packaging of the adenovirus genes into the capsid. Alternatively, one or more of these proteins (whether hybrid or non-hybrid) can be functionally deleted in the
30 chimeric adenovirus. Where desired, any necessary proteins functionally deleted in the chimeric adenovirus can be expressed in *trans* in the packaging cell.

B. Parental Adenovirus Structural Proteins

1. Serotypes

This invention is particularly well adapted for use in generating chimeric adenoviruses in which the capsid proteins are from a parental
5 adenovirus which does not efficiently grow in a desirable host cell. The selection of the parental adenovirus serotype providing the internal regions can be readily made by one of skill in the art based on the information provided herein.

A variety of suitable adenoviruses can serve as a parental adenovirus supplying the regions encoding the structural (*i.e.*, capsid proteins). Many
10 of such adenoviruses can be obtained from the same sources as described above for the donating adenovirus serotypes. Examples of suitable parental adenovirus serotypes includes, without limitation, human adenovirus serotype 40, among others [see, *e.g.*, Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)], and adenoviruses known to infect non-human primates (*e.g.*,
15 chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals, including, without limitation, chimpanzee adenovirus C1, described in US Patent No. 6,083,716, which is incorporated by reference; simian adenoviruses, and baboon adenoviruses, among others. In addition, the parental adenovirus supplying the internal regions may be from a non-naturally occurring adenovirus serotype, such
20 as may be generated using a variety of techniques known to those of skill in the art.

In one embodiment illustrated herein, a chimeric virus that was constructed was that between the chimpanzee adenoviruses Pan-5 and C1 exhibited a higher titer in human 293 cells than the wild-type parental virus. However, the invention is not limited to the use of these chimpanzee adenoviruses, or to the combination of simian-
25 simian, human-human, or simian-human chimeric adenoviruses. For example, it may be desirable to utilize bovine or canine adenoviruses, or other non-human mammalian adenoviruses which do not naturally infect and/or replicate in human cells.

In the following examples, the human adenovirus type 40 (Ad40) and the chimpanzee adenovirus C1, simian Pan 5 and Ad40, and Pan 5 and
30 simian adenovirus SA18, are used. However, one of skill in the art will understand that other adenoviral serotypes may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The parental adenovirus provides to the chimeric construct of the invention its internal regions which includes structural proteins necessary for generating a capsid having the desired characteristics of the parental adenovirus.

5 These desired characteristics include, but are not limited to, the ability to infect target cells and delivery a heterologous transgene, the ability to elude neutralizing antibodies directed to another adenovirus serotype (*i.e.*, avoiding clearance due to cross-reactivity), and/or the ability to infect cells in the absence of an immune response to the chimeric adenovirus. The advantages of such characteristics may be

10 most readily apparent in a regimen which involves repeat delivery of adenoviral vectors. The left and right termini of the parent adenovirus, including at least the 5' ITRs, the E1 region, the E4 region and the 3' ITRs are non-functional and, preferably, completely absent. Optionally, all adenovirus regulatory proteins from this parental

15 proteins) are retained.

At a minimum, the parental adenovirus provides the adenoviral late region encoding the hexon protein. Suitably, the parental adenovirus further provides the late regions encoding the penton and the fiber. In certain embodiments, all of the functional adenoviral late regions, including L1 (encoding

20 52/55 Da, IIIa proteins), L2 (encoding penton, VII, V, Mu proteins), L3 (encoding VI, hexon, Endoprotease), L4 (encoding 100 kD, 33 kD, VIII proteins) and L5 (encoding fiber protein) are supplied by the parental adenovirus. Optionally, one or more of these late gene functions, with the exception of those encoding the hexon, penton and fiber proteins, can be functionally deleted. Any necessary structural proteins may be

25 supplied in *trans*.

Thus, in certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the

30 parental adenovirus donating its internal regions. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins, as described above.

C. The "Minigene"

Typically, an adenoviral vector of the invention is designed to contain a minigene which may be inserted into the site of a partially deleted, fully deleted (absent), or disrupted adenoviral gene. For example, the minigene may be located in
5 the site of such a functional E1 deletion or functional E3 deletion, or another suitable site.

The methods employed for the selection of the transgene, the cloning and construction of the "minigene" and its insertion into the viral vector are within the skill in the art given the teachings provided herein.

10 1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation,
15 and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the adenoviral vector will be put. For example, the adenoviral vector may be used as a helper virus in production of recombinant adeno-associated viruses or in production of recombinant adenoviruses deleted of essential adenoviral gene
20 functions which are supplied by the adenoviral vector, or for a variety of production uses. Alternatively, the adenoviral vector may be used for diagnostic purposes.

One type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase
25 (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane
30 bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other

spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for
5 beta-galactosidase activity. Where the transgene is GFP or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins,
10 peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, si RNAs, small hairpin RNAs, trans-splicing RNAs, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

15 The transgene may be used for treatment, *e.g.*, of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a molecule (*e.g.*, a gene product) to induce a T cell and/or a humoral immune response to the molecule. The invention further includes
20 using multiple transgenes, *e.g.*, to correct or ameliorate a condition caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, *e.g.*, for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order
25 for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when
30 the size of the DNA encoding each of the subunits is small, *e.g.*, the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, *e.g.*, M.L. Donnelly, *et al*, *J. Gen.*

Virol., 78(Pt 1):13-21 (Jan 1997); Furler, S., *et al*, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, 8(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, *e.g.*, a product desirable for study.

Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Vector and Transgene Regulatory Elements

In addition to the major elements identified above for the minigene, the adenoviral vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*i.e.*, Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, *e.g.*, Boshart *et al*, *Cell*, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter [Invitrogen].

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as

temperature, or the presence of a specific physiological state, *e.g.*, acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems

5 have been described and can be readily selected by one of skill in the art. For example, inducible promoters include the zinc-inducible sheep metallothioneine (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, **93**:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*, *Proc. Natl. Acad. Sci. USA*, **89**:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, **268**:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, **2**:512-518 (1998)]. Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system [Wang *et al*, *Nat. Biotech.*, **15**:239-243 (1997) and Wang *et al*, *Gene Ther.*, **4**:432-441 (1997)] and the rapamycin-inducible system [Magari *et al*, *J. Clin. Invest.*, **100**:2865-2872 (1997)]. The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, *e.g.*, TetR linked to a TetR by an IRES. Alternatively, one can

20 wait at least 3 days before screening for the desired function. One can enhance expression of desired proteins by known means to enhance the effectiveness of this system. For example, using the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE).

In another embodiment, the native promoter for the transgene

25 will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control

30 elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal

muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li *et al.*, *Nat. Biotech.*, 17:241-245 (1999)).

5 Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.*, *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, 7:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, 24:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, 11:654-64

10 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli *et al.*, *Neuron*, 15:373-84 (1995)),

15 among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes

20 (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, *e.g.*, Sambrook *et al.*, and references cited

25 therein].

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*,

30 Cold Spring Harbor Press, Cold Spring Harbor, NY], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

II. Production of the Recombinant Viral Particle

In one embodiment, the invention provides a method of generating recombinant chimeric adenoviral particles in which the capsid of the chimeric adenovirus is of a serotype incapable of efficient growth in the selected host cell. A
5 vector suitable for production of recombinant chimeric adenoviral particles can be generated by direct cloning. Alternatively, such particles can be generated by homologous recombination between a first vector containing the left end of the chimeric adenoviral genome and a second vector containing the right end of the chimeric adenoviral genome. However, any suitable methodology known to those of
10 skill in the art can be readily utilized to generate a vector suitable to generate a production vector, preferably which contains the entire chimeric adenoviral genome, including a minigene. This production vector is then introduced into a host cell in which the adenoviral capsid protein is assembled and the chimeric adenoviral particle assembled as described.

15 The chimeric adenoviruses of the invention include those in which one or more adenoviral genes are absent, or otherwise rendered non-functional. If any of the missing gene functions are essential to the replication and infectivity of the adenoviral particle, these functions are supplied by a complementation (or transcomplementing) cell line or a helper vector expressing these functions during production of the
20 chimeric adenoviral particle.

Examples of chimeric adenoviruses containing such missing adenoviral gene functions include those which are partially or completely deleted in the E1a and/or E1b gene. In such a case, the E1 gene functions can be supplied by the packaging host cell, permitting the chimeric construct to be deleted of E1 gene functions and, if
25 desired, for a transgene to be inserted in this region. Optionally, the E1 gene can be of a serotype which transcomplements the serotype providing the other adenovirus sequences in order to further reduce the possibility of recombination and improve safety. In other embodiments, it is desirable to retain an intact E1a and/or E1b region in the recombinant adenoviruses. Such an intact E1 region may be located in its
30 native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (*e.g.*, in the E3 region).

In another example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the chimeric adenovirus. The function of adenovirus E3 is

believed to be irrelevant to the function and production of the recombinant virus particle. Chimeric adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of this invention contains a deletion in the delayed early gene E2a. Similarly, deletions in the intermediate genes IX and IVa₂ may be useful for some purposes. Optionally, deletions may also be made in selected portions of the late genes L1 through L5, as described above.

Other deletions may be made in the other structural or non-structural adenovirus genes. The above-discussed deletions may be used individually, *i.e.*, an adenovirus sequence for use in the present invention may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such deletions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

Examples of suitable transcomplementing serotypes are provided above. The use of transcomplementing serotypes can be particularly advantageous where there is diversity between the Ad sequences in the vector of the invention and the human AdE1 sequences found in currently available packaging cells. In such cases, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process. However, in certain circumstances, it will be desirable to utilize a cell line which expresses the E1 gene products can be utilized for production of an E1-deleted simian adenovirus. Such cell lines have been described. See, *e.g.*, US Patent 6,083,716.

A. Packaging Host Cells

Suitably, the packaging host cell is selected from among cells in which the adenovirus serotype donating the left and right terminal ends of the chimeric genome are capable of efficient growth. The host cells are preferably of mammalian origin, and most preferably are of non-human primate or human origin.

Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549 [ATCC Accession No. CCL 185], 911 cells, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ *et al*, (1998), *Hum Gene Ther*, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2, HeLa [ATCC Accession No. CCL 2], KB [CCL 17], Detroit [*e.g.*, Detroit 510, CCL 72] and WI-38 [CCL 75] cells, and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209. Other suitable cell lines may be obtained from other sources. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, *i.e.*, fibroblast, hepatocyte, tumor cell, etc.

As described above, a chimeric adenovirus of the invention can lack one or more functional adenoviral regulatory and/or structural genes which are supplied either by the host cell or in *trans* to effect packaging of the chimeric adenovirus into the viral capsid to generate the viral particle. Thus, the ability of a selected host cell to supply transcomplementing adenoviral sequences may be taken into consideration in selecting a desired host cell.

In one example, the cells are from a stable cell line which expresses adenovirus E1a and E1b functions from a cell line which transcomplements the adenovirus serotype which donates the left and right termini to the chimera of the invention, permitting the chimera to be E1-deleted. Alternatively, where the cell line does not transcomplement the adenovirus donating the termini, E1 functions may be provided by the chimera, or in *trans*.

If desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the adenovirus E1 gene from the adenovirus serotype donating the 5' ITR under the transcriptional control of a promoter for expression, or a transcomplementing serotype, in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this specification. A parent cell is selected for the generation of a novel cell line expressing any desired adenovirus or adenovirus gene, including, *e.g.*, a human Ad5, AdPan5, Pan6, Pan7,

SV1, SV25 or SV39 gene. Without limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [*e.g.*, Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. Many of these cell lines are all available from the ATCC. Other suitable
5 parent cell lines may be obtained from other sources.

Such E1-expressing cell lines are useful in the generation of chimeric adenovirus E1 deleted vectors. Additionally, or alternatively, the invention provides cell lines that express one or more simian adenoviral gene products, *e.g.*, E1a, E1b, E2a, and/or E4 ORF6, can be constructed using essentially the same procedures for
10 use in the generation of chimeric viral vectors. Such cell lines can be utilized to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (*e.g.*, adeno-associated virus). The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA
15 sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide sequences of the adenovirus genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired
20 nucleotide sequence.

In still another alternative, the essential adenoviral gene products are provided in *trans* by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (*e.g.*, bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and
25 mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from
30 mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, *i.e.*, fibroblast, hepatocyte, tumor cell, etc.

B. Helper Vectors

Thus, depending upon the adenovirus gene content of the adenoviral vectors and any adenoviral gene functions expressed from the host cell, a helper vector may be necessary to provide sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. See, for example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597, published May 9, 1996, and incorporated herein by reference. Suitably, these helper vectors may be non-replicating genetic elements, a plasmid, or a virus.

Useful helper vectors contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains a variety of adenovirus genes in addition to the sequences described above. Such a helper vector is desirably used in combination with an E1-expressing cell line.

Helper vectors may be formed into poly-cation conjugates as described in Wu *et al*, *J. Biol. Chem.*, **264**:16985-16987 (1989); K. J. Fisher and J. M. Wilson, *Biochem. J.*, **299**:49 (April 1, 1994). A helper vector may optionally contain a second reporter minigene. A number of such reporter genes are known to the art. The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the Ad vector and the helper vector to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

C. Assembly of Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 μ g to about 100 μ g DNA, and preferably about 10 to about 50 μ g DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the

host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Furthermore, the promoters for each of the adenoviral genes may be
5 selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (*i.e.*, by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

Introduction of the molecules (as plasmids or viruses) into the host
10 cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, *e.g.*, CaPO₄ transfection or electroporation.

Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and
15 the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include direct cloning as described [G. Gao *et al*, *Gene Ther.* 2003 Oct; **10**(22):1926-1930; US Patent Publication No. 2003-0092161-A, published May 15, 2003; International Patent Application No. PCT/US03/12405]. Other cloning techniques of cDNA such as
20 those described in texts [Sambrook *et al*, cited above], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence can be utilized. Standard transfection and co-transfection techniques are employed, *e.g.*, CaPO₄ precipitation techniques. Other conventional methods employed include homologous
25 recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected *in vitro* in the presence of an optional helper vector into the packaging cell line. The functions expressed from
30 the plasmid, packaging cell line and helper virus, if any, permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the chimeric viral particles. The current method for producing such virus particles is transfection-based. However, the invention is not limited to such

methods. The resulting chimeric adenoviruses are useful in transferring a selected transgene to a selected cell.

III. Use of the Chimeric Adenovirus Vectors

5 The chimeric adenovirus vectors of the invention are useful for gene transfer to a human or veterinary subject (including, non-human primates, non-simian primates, and other mammals) *in vitro*, *ex vivo*, and *in vivo*.

The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous genes *in vitro*. For example, the recombinant adenoviruses containing a gene inserted
10 into the location of an E1 deletion may be transfected into an E1-expressing cell line as described above. Alternatively, replication-competent adenoviruses may be used in another selected cell line. The transfected cells are then cultured in the conventional manner, allowing the recombinant adenovirus to express the gene
15 product from the promoter. The gene product may then be recovered from the culture medium by known conventional methods of protein isolation and recovery from culture.

A chimeric adenoviral vector of the invention provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell *in vivo* or
20 *ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAd and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective
25 immunity.

More commonly, the chimeric adenoviral vectors of the invention will be utilized for delivery of therapeutic or immunogenic molecules, as described below. It will be readily understood for both applications that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving
30 repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (*e.g.*, one, two, three, four

or more). Thus, a regimen may involve delivery of a rAd with a first capsid, delivery with a rAd with a second capsid, and delivery with a third capsid. A variety of other regimens which use the Ad capsids of the invention alone, in combination with one another, or in combination with other Ad serotypes will be apparent to those of skill
5 in the art. Optionally, such a regimen may involve administration of rAd with capsids of non-human primate adenoviruses, human adenoviruses, or artificial (*e.g.*, chimeric) serotypes such as are described herein. Each phase of the regimen may involve administration of a series of injections (or other delivery routes) with a single Ad serotype capsid followed by a series with another Ad serotype capsid. Alternatively,
10 the recombinant Ad vectors of the invention may be utilized in regimens involving other non-adenoviral-mediated delivery systems, including other viral systems, non-viral delivery systems, protein, peptides, and other biologically active molecules.

The following sections will focus on exemplary molecules which may be delivered via the adenoviral vectors of the invention.

15 A. Ad-Mediated Delivery of Therapeutic Molecules

In one embodiment, the Ad vectors described herein are administered to humans according to published methods for gene therapy. A viral vector of the invention bearing the selected transgene may be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable
20 delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

The adenoviral vectors are administered in sufficient amounts to
25 transduce the target cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the retina and other intraocular
30 delivery methods, direct delivery to the liver, inhalation, intranasal, intravenous, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parenteral routes of administration. Routes of administration may be combined, if

desired, or adjusted depending upon the transgene or the condition. The route of administration primarily will depend on the nature of the condition being treated.

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector is generally in the range of from about 100 μ L to about 100 mL of a carrier containing concentrations of from about 1×10^6 to about 1×10^{15} particles, about 1×10^{11} to 1×10^{13} particles, or about 1×10^9 to 1×10^{12} particles. Dosages will range depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be delivered. In another example, a suitable human or veterinary dosage may be in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation. One of skill in the art may adjust these doses, depending the route of administration, and the therapeutic or vaccinal application for which the recombinant vector is employed. The levels of expression of the transgene, or for an immunogen, the level of circulating antibody, can be monitored to determine the frequency of dosage administration. Yet other methods for determining the timing of frequency of administration will be readily apparent to one of skill in the art.

An optional method step involves the co-administration to the patient, either concurrently with, or before or after administration of the viral vector, of a suitable amount of a short acting immune modulator. The selected immune modulator is defined herein as an agent capable of inhibiting the formation of neutralizing antibodies directed against the recombinant vector of this invention or capable of inhibiting cytolytic T lymphocyte (CTL) elimination of the vector. The immune modulator may interfere with the interactions between the T helper subsets (T_{H1} or T_{H2}) and B cells to inhibit neutralizing antibody formation. Alternatively, the immune modulator may inhibit the interaction between T_{H1} cells and CTLs to reduce the occurrence of CTL elimination of the vector. A variety of useful immune modulators and dosages for use of same are disclosed, for example, in Yang *et al.*, *J. Virol.*, 70(9) (Sept 1996); International Patent Application No. WO96/12406, published May 2, 1996; and International Patent Application No. PCT/US96/03035,

all incorporated herein by reference. Typically, such immune modulators would be selected when the transgene is a therapeutic which requires repeat delivery.

1. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include

5 hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF),
10 erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor α (TGF α), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor superfamily, including TGF, activins, inhibins, or any of
15 the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of
20 semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, *e.g.*, IL-2, IL-
25 4, IL-12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors and, interferons, and, stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitation, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins,
30 humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include

complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, proteins useful in the regulation of lipids, including, *e.g.*, apolipoprotein (apo) A and its isoforms (*e.g.*, ApoAI), apoE and its isoforms including E2, E3 and E4), SRB1, ABC1, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as jun, fos, max, mad, serum response factor (SRF), AP-1, AP2, myb, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, *e.g.*, GATA-3, and the forkhead family of winged helix proteins.

Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetic acid hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylase, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Other useful gene products include those useful for treatment of hemophilia A (*e.g.*, Factor VIII and its variants, including the light chain and heavy chain of the heterodimer, optionally operably linked by a junction), and the B-domain deleted Factor VIII, see US 6,200,560 and 6,221,349], and useful for treatment of hemophilia B (*e.g.*, Factor IX).

Still other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful
5 in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene are particularly desirable for treatment of hyperproliferative conditions characterized by
10 hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGFR. In addition to oncogene products as
15 target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found
20 at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that
25 are associated with autoimmunity including cell receptors and cells which produce self-directed antibodies. T-cell mediated autoimmune diseases include rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis,
30 vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

The chimeric adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes is desired, *e.g.*, in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a chimeric adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a chimeric adenoviral vector of the invention, in which the serotype of the viral vector delivered in the first administration differs from the serotype of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a chimeric vector and repeat administration with one or more adenoviral vectors of the same or different serotypes. In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a chimeric vector of the invention which differs from the serotype of the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the serotype of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the chimeric serotypes of the invention. Rather, these regimens can readily utilize chimeric or non-chimeric vectors of other adenoviral serotypes, which may be of artificial, human or non-human primate, or other mammalian sources, in combination with one or more of the chimeric vectors of the invention. Examples of such serotypes are discussed elsewhere in this document. Further, these therapeutic regimens may involve either simultaneous or sequential delivery of chimeric adenoviral vectors of the invention in combination with non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules. The present invention is not limited to these therapeutic regimens, a variety of which will be readily apparent to one of skill in the art.

B. Ad-Mediated Delivery of Immunogenic Transgenes

The adenoviruses of the invention may also be employed as immunogenic compositions. As used herein, an immunogenic composition is a composition to which a humoral (*e.g.*, antibody) or cellular (*e.g.*, a cytotoxic T cell) response is mounted to a transgene product delivered by the immunogenic composition following delivery to a mammal, and preferably a primate. The present

invention provides an Ad that can contain in any of its adenovirus sequence deletions a gene encoding a desired immunogen. Chimeric adenoviruses based on simian or other non-human mammalian primate serotypes are likely to be better suited for use as a live recombinant virus vaccine in different animal species compared to an
5 adenovirus of human origin, but is not limited to such a use. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an immune response and able to limit the spread of the pathogen has been identified and for which the cDNA is available.

10 Such vaccinal (or other immunogenic) compositions are formulated in a suitable delivery vehicle, as described above. Generally, doses for the immunogenic compositions are in the range defined above for therapeutic compositions. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in
15 the serum, optional booster immunizations may be desired.

Optionally, a vaccinal composition of the invention may be formulated to contain other components, including, *e.g.* adjuvants, stabilizers, pH adjusters, preservatives and the like. Such components are well known to those of skill in the vaccine art. Examples of suitable adjuvants include, without limitation, liposomes,
20 alum, monophosphoryl lipid A, and any biologically active factor, such as cytokine, an interleukin, a chemokine, a ligands, and optimally combinations thereof. Certain of these biologically active factors can be expressed *in vivo*, *e.g.*, via a plasmid or viral vector. For example, such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response
25 compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only.

The adenoviruses are administered in "an immunogenic amount", that is, an amount of adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to
30 induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

Alternatively, or in addition, the vectors of the invention may contain a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. The recombinant adenoviruses of this invention are expected to be highly efficacious at inducing cytolytic T cells and antibodies to the inserted heterologous antigenic protein expressed by the vector.

For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. In addition, the human coronaviruses include the putative causative agent of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (*e.g.*, Vesicular

Stomatitis Virus), and the general lyssavirus (*e.g.*, rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus, may be a suitable source of antigens. The paramyxovirus family
5 includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus), parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the
10 family orthomyxovirus and is a suitable source of antigen (*e.g.*, the HA protein, the N1 protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (puremala is a hemahagin fever virus), nairovirus (Nairobi sheep disease) and various unassigned bungaviruses. The arenavirus family provides a source of antigens against LCM and
15 Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever), Lebombo (humans), equine encephalosis, blue tongue.

The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I
20 and HTLV II, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Among the lentiviruses, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat,
25 Nef, and Rev proteins, as well as various fragments thereof. For example, suitable fragments of the Env protein may include any of its subunits such as the gp120, gp160, gp41, or smaller fragments thereof, *e.g.*, of at least about 8 amino acids in length. Similarly, fragments of the tat protein may be selected. [See, US Patent 5,891,994 and US Patent 6,193,981.] See, also, the HIV and SIV proteins described
30 in D.H. Barouch *et al*, *J. Virol.*, 75(5):2462-2467 (March 2001), and R.R. Amara, *et al*, *Science*, 292:69-74 (6 April 2001). In another example, the HIV and/or SIV immunogenic proteins or peptides may be used to form fusion proteins or other immunogenic molecules. See, *e.g.*, the HIV-1 Tat and/or Nef fusion proteins and

immunization regimens described in WO 01/54719, published August 2, 2001, and WO 99/16884, published April 8, 1999. The invention is not limited to the HIV and/or SIV immunogenic proteins or peptides described herein. In addition, a variety of modifications to these proteins have been described or could readily be made by one of skill in the art. See, *e.g.*, the modified gag protein that is described in US Patent 5,972,596. Further, any desired HIV and/or SIV immunogens may be delivered alone or in combination. Such combinations may include expression from a single vector or from multiple vectors. Optionally, another combination may involve delivery of one or more expressed immunogens with delivery of one or more of the immunogens in protein form. Such combinations are discussed in more detail below.

The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus includes family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSVI, HSVII), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitts lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

The viruses of the present invention may also carry immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell.

Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; 5 salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella; *Francisella tularensis* (which causes tularemia); yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelotheirus rhusiopathiae; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and 10 bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, 15 histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; 20 lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; Trichans; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes 25 or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), 30 *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (*e.g.*, Ebola, Marburg), and arenaviruses (*e.g.*, Lassa, Machupo)], all of which are currently classified as Category A agents; *Coxiella burnetii* (Q fever);

Brucella species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (meloidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), Staphylococcus species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Rickettsia powazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In RA, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V α -17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In MS, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and V α -10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V α -16, V α -3C, V α -7, V α -14, V α -15, V α -16, V α -28 and V α -12. Thus, delivery of a chimeric adenovirus that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

C. Ad-Mediated Delivery Methods

The therapeutic levels, or levels of immunity, of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster

immunizations may be desired. Optionally, the adenoviral vectors of the invention may be delivered in a single administration or in various combination regimens, *e.g.*, in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen. A variety of such regimens have been
5 described in the art and may be readily selected.

For example, prime-boost regimens may involve the administration of a DNA (*e.g.*, plasmid) based vector to prime the immune system to a second or further, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. See, *e.g.*, WO
10 00/11140, published March 2, 2000, incorporated by reference. Alternatively, an immunization regimen may involve the administration of a chimeric adenoviral vector of the invention to boost the immune response to a vector (either viral or DNA-based) carrying an antigen, or a protein. In still another alternative, an immunization regimen involves administration of a protein followed by booster with a vector
15 encoding the antigen.

In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting with an adenoviral vector of the invention. In one embodiment, the prime-boost regimen involves the expression of
20 multiproteins from the prime and/or the boost vehicle. See, *e.g.*, R.R. Amara, *Science*, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and
25 HIV-1 Env is delivered in a recombinant adenovirus construct of the invention. Still other regimens are described in WO 99/16884 and WO 01/54719.

However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first vector of the invention followed by boosting with a second
30 vector, or with a composition containing the antigen itself in protein form. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can

be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

The priming composition may be administered at various sites in the body in a dose dependent manner, which depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or situs of injection(s) or to the pharmaceutical carrier. Rather, the regimen may involve a priming and/or boosting step, each of which may include a single dose or dosage that is administered hourly, daily, weekly or monthly, or yearly. As an example, the mammals may receive one or two doses containing between about 10 µg to about 50 µg of plasmid in carrier. A desirable amount of a DNA composition ranges between about 1 µg to about 10,000 µg of the DNA vector. Dosages may vary from about 1 µg to 1000 µg DNA per kg of subject body weight. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal. The dosage unit of the vector suitable for delivery of the antigen to the mammal is described herein. The vector is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes. Optionally, the priming step of this invention also includes administering with the priming composition, a suitable amount of an adjuvant, such as are defined herein.

Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source (*e.g.*, adenoviral sequences of the invention) or from another source. Alternatively, the “boosting composition” can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which

composition induces an immune response in the host. In another embodiment, the boosting composition contains a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, *e.g.*, vectors such as well-known bacterial or viral vectors. The primary requirements of
5 the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

In another embodiment, the adenoviral vectors of the invention are also well suited for use in a variety of other immunization and therapeutic regimens. Such regimens may involve delivery of adenoviral vectors of the invention
10 simultaneously or sequentially with Ad vectors of different serotype capsids, regimens in which adenoviral vectors of the invention are delivered simultaneously or sequentially with non-Ad vectors, regimens in which the adenoviral vectors of the invention are delivered simultaneously or sequentially with proteins, peptides, and/or other biologically useful therapeutic or immunogenic compounds. Such uses will be
15 readily apparent to one of skill in the art.

IV. Simian Adenovirus 18 Sequences

The invention provides nucleic acid sequences and amino acid sequences of Ad SA18, which are isolated from the other viral material with which they are
20 associated in nature. These sequences are useful in preparing heterologous molecules containing the nucleic acid sequences and amino acid sequences, and regions or fragments thereof as are described herein, viral vectors which are useful for a variety of purposes, including the constructs and compositions, and such methods as are described herein for the chimeric adenoviruses, including, *e.g.*, in host cells for
25 production of viruses requiring adenoviral helper functions, as delivery vehicles for heterologous molecules such as those described herein. These sequences are also useful in generating the chimeric adenoviruses of the invention.

A. Nucleic Acid Sequences

The SA18 nucleic acid sequences of the invention include nucleotides
30 SEQ ID NO: 12, nt 1 to 31967. See, Sequence Listing, which is incorporated by reference herein. The nucleic acid sequences of the invention further encompass the strand which is complementary to the sequences of SEQ ID NO: 12, as well as the RNA and cDNA sequences corresponding to the sequences of these sequences figures

and their complementary strands. Further included in this invention are nucleic acid sequences which are greater than 95 to 98%, and more preferably about 99 to 99.9% homologous or identical to the Sequence Listing. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the sequences provided in SEQ ID NO: 12 and their complementary strands. Such modifications include, for example, labels that are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

The invention further encompasses fragments of the sequences of SA18, their complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, *i.e.*, fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of recombinant viral vectors. Such fragments include the gene sequences and fragments listed in the tables below.

The following tables provide the transcript regions and open reading frames in the simian adenovirus sequences of the invention. For certain genes, the transcripts and open reading frames (ORFs) are located on the strand complementary to that presented in SEQ ID NO: 12. See, *e.g.*, E2b, E4 and E2a. The calculated molecular weights of the encoded proteins are also shown.

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
		start	End	M.W.
ITR		1	180	
E1a	13S	916	1765	27264
	12S	916	1765	24081
E1b	Small T	1874	2380	19423
	LargeT	2179	3609	52741
	IX	3678	4079	13701
E2b	IVa2	5478	4126	51295
	Polymerase	13745	5229	128392
	PTP	13745	8597	75358
	Agnoprotein	8007	8705	23610
L1	52/55 kD	10788	11945	43416
	IIIa	11966	13699	63999
L2	Penton	13796	15322	57166
	VII	15328	15873	20352
	V	15920	17050	42020
L3	VI	17348	18154	29222
	Hexon	18257	21010	102912
	Endoprotease	21029	21640	23015

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
2a	DBP	23147	21711	53626
L4	100kD	23175	25541	87538
	22 kD homolog	25204	25797	22206
	33 kD homolog	25204	26025	24263
	VIII	26107	26817	25490
E3	Orf #1	26817	27125	11814
L5	Fiber	27192	29015	65455
E4	Orf 6/7	30169	29067	13768
	Orf 6	30169	29303	33832
	Orf 4	30464	30099	14154
	Orf 3	30816	30466	13493
	Orf 2	31205	30813	14698
	Orf 1	31608	31231	14054
ITR		31788	31967	

The SA18 adenoviral nucleic acid sequences are useful as therapeutic and immunogenic agents and in construction of a variety of vector systems and host cells.

- 5 Such vectors are useful for any of the purposes described above for the chimeric adenovirus. Additionally, these SA18 sequences and products may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from other adenoviral or non-adenoviral sequences. The adenoviral sequences of the invention are also useful as antisense delivery vectors, gene therapy
- 10 vectors, or vaccine vectors, and in methods of using same. Thus, the invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the Ad sequences of the invention.

- For example, the invention encompasses a nucleic acid molecule containing simian Ad ITR sequences of the invention. In another example, the invention
- 15 provides a nucleic acid molecule containing simian Ad sequences of the invention

encoding a desired Ad gene product. Still other nucleic acid molecule constructed using the sequences of the invention will be readily apparent to one of skill in the art, in view of the information provided herein.

In one embodiment, the simian Ad gene regions identified herein may be used
5 in a variety of vectors for delivery of a heterologous molecule to a cell. Examples of such molecules and methods of delivery are provided in Section III herein. For example, vectors are generated for expression of an adenoviral capsid protein (or fragment thereof) for purposes of generating a viral vector in a packaging host cell. Such vectors may be designed for expression in trans. Alternatively, such vectors are
10 designed to provide cells which stably contain sequences which express desired adenoviral functions, *e.g.*, one or more of E1a, E1b, the terminal repeat sequences, E2a, E2b, E4, E4ORF6 region.

In addition, the adenoviral gene sequences and fragments thereof are useful for providing the helper functions necessary for production of helper-dependent
15 viruses (*e.g.*, adenoviral vectors deleted of essential functions or adeno-associated viruses (AAV)). For such production methods, the simian adenoviral sequences of the invention are utilized in such a method in a manner similar to those described for the human Ad. However, due to the differences in sequences between the simian adenoviral sequences of the invention and those of human Ad, the use of the
20 sequences of the invention essentially eliminate the possibility of homologous recombination with helper functions in a host cell carrying human Ad E1 functions, *e.g.*, 293 cells, which may produce infectious adenoviral contaminants during rAAV production.

Methods of producing rAAV using adenoviral helper functions have been
25 described at length in the literature with human adenoviral serotypes. See, *e.g.*, US Patent 6,258,595 and the references cited therein. See, also, US Patent 5,871,982; WO 99/14354; WO 99/15685; WO 99/47691. These methods may also be used in production of non-human serotype AAV, including non-human primate AAV serotypes. The simian adenoviral gene sequences of the invention which provide the
30 necessary helper functions (*e.g.*, E1a, E1b, E2a and/or E4 ORF6) can be particularly useful in providing the necessary adenoviral function while minimizing or eliminating the possibility of recombination with any other adenoviruses present in the rAAV-packaging cell which are typically of human origin. Thus, selected genes or open

reading frames of the adenoviral sequences of the invention may be utilized in these rAAV production methods.

Alternatively, recombinant adenoviral simian vectors of the invention may be utilized in these methods. Such recombinant adenoviral simian vectors may include, *e.g.*, a hybrid simian Ad/AAV in which simian Ad sequences flank a rAAV expression cassette composed of, *e.g.*, AAV 3' and/or 5' ITRs and a transgene under the control of regulatory sequences which control its expression. One of skill in the art will recognize that still other simian adenoviral vectors and/or gene sequences of the invention will be useful for production of rAAV and other viruses dependent upon adenoviral helper.

In still another embodiment, nucleic acid molecules are designed for delivery and expression of selected adenoviral gene products in a host cell to achieve a desired physiologic effect. For example, a nucleic acid molecule containing sequences encoding an adenovirus E1a protein of the invention may be delivered to a subject for use as a cancer therapeutic. Optionally, such a molecule is formulated in a lipid-based carrier and preferentially targets cancer cells. Such a formulation may be combined with other cancer therapeutics (*e.g.*, cisplatin, taxol, or the like). Still other uses for the adenoviral sequences provided herein will be readily apparent to one of skill in the art.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules, including any of those identified as being deliverable via the chimeric adenoviruses of the invention. For example, the simian Ad genome of the invention can be utilized in a variety of rAd and non-rAd vector systems. Such vectors systems may include, *e.g.*, plasmids, lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adeno-associated viral systems, among others. Selection of these vector systems is not a limitation of the present invention.

The invention further provides molecules useful for production of the simian and simian-derived proteins of the invention. Such molecules which carry polynucleotides including the simian Ad DNA sequences of the invention can be in the form of a vector.

B. Simian Adenoviral Proteins of the Invention

The invention further provides gene products of the above adenoviruses, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids of the invention. The invention further encompasses
5 SA18 proteins, enzymes, and fragments thereof, having the amino acid sequences encoded by these nucleic acid sequences which are generated by other methods. Such proteins include those encoded by the open reading frames identified in the tables above, and fragments thereof.

Thus, in one aspect, the invention provides unique simian adenoviral
10 proteins which are substantially pure, *i.e.*, are free of other viral and proteinaceous proteins. Preferably, these proteins are at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

In one embodiment, the invention provides unique simian-derived capsid proteins. As used herein, a simian-derived capsid protein includes any
15 adenoviral capsid protein that contains a SA18 capsid protein or a fragment thereof, as defined above, including, without limitation, chimeric capsid proteins, fusion proteins, artificial capsid proteins, synthetic capsid proteins, and recombinantly capsid proteins, without limitation to means of generating these proteins.

Suitably, these simian-derived capsid proteins contain one or more
20 SA18 regions or fragments thereof (*e.g.*, a hexon, penton, fiber or fragment thereof) in combination with capsid regions or fragments thereof of different adenoviral serotypes, or modified simian capsid proteins or fragments, as described herein. A "modification of a capsid protein associated with altered tropism" as used herein includes an altered capsid protein, *i.e.*, a penton, hexon or fiber protein region, or
25 fragment thereof, such as the knob domain of the fiber region, or a polynucleotide encoding same, such that specificity is altered. The simian-derived capsid may be constructed with one or more of the simian Ad of the invention or another Ad serotypes which may be of human or non-human origin. Such Ad may be obtained from a variety of sources including the ATCC, commercial and academic sources, or
30 the sequences of the Ad may be obtained from GenBank or other suitable sources.

The amino acid sequences of the simian adenoviruses penton proteins of the invention are provided herein. The AdSA18 penton protein is provided in SEQ ID NO: 13. Suitably, any of these penton proteins, or unique fragments thereof, may

be utilized for a variety of purposes. Examples of suitable fragments include the penton having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. Further, the
5 penton protein may be modified for a variety of purposes known to those of skill in the art.

The invention further provides the amino acid sequences of the hexon protein of SA18, SEQ ID NO:14. Suitably, this hexon protein, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments
10 include the hexon having N-terminal and/or C-terminal truncations of about 50, 100, 150, 200, 300, 400, or 500 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 14. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. For example, one suitable fragment the loop region (domain) of the hexon protein, designated DE1 and FG1, or a
15 hypervariable region thereof. Such fragments include the regions spanning amino acid residues about 125 to 443; about 138 to 441, or smaller fragments, such as those spanning about residue 138 to residue 163; about 170 to about 176; about 195 to about 203; about 233 to about 246; about 253 to about 264; about 287 to about 297; about 404 to about 430, about 430 to 550, about 545 to 650; of the simian hexon
20 proteins, with reference to SEQ ID NO: 14. Other suitable fragments may be readily identified by one of skill in the art. Further, the hexon protein may be modified for a variety of purposes known to those of skill in the art. Because the hexon protein is the determinant for serotype of an adenovirus, such artificial hexon proteins would result in adenoviruses having artificial serotypes. Other artificial capsid proteins can
25 also be constructed using the chimp Ad penton sequences and/or fiber sequences of the invention and/or fragments thereof.

In one example, it may be desirable to generate an adenovirus having an altered hexon protein utilizing the sequences of a hexon protein of the invention. One suitable method for altering hexon proteins is described in US Patent 5,922,315,
30 which is incorporated by reference. In this method, at least one loop region of the adenovirus hexon is changed with at least one loop region of another adenovirus serotype. Thus, at least one loop region of such an altered adenovirus hexon protein is a simian Ad hexon loop region of the invention. In one embodiment, a loop region

of the SA18 hexon protein is replaced by a loop region from another adenovirus serotype. In another embodiment, the loop region of the SA18 hexon is used to replace a loop region from another adenovirus serotype. Suitable adenovirus serotypes may be readily selected from among human and non-human serotypes, as
5 described herein. SA18 is selected for purposes of illustration only; the other simian Ad hexon proteins of the invention may be similarly altered, or used to alter another Ad hexon. The selection of a suitable serotype is not a limitation of the present invention. Still other uses for the hexon protein sequences of the invention will be readily apparent to those of skill in the art.

10 The invention further encompasses the fiber proteins of the simian adenoviruses of the invention. The fiber protein of AdSA18 has the amino acid sequence of SEQ ID NO: 15. Suitably, this fiber protein, or unique fragments thereof, may be utilized for a variety of purposes. One suitable fragment is the fiber knob, which spans about amino acids 247 to 425 of SEQ ID NO: 15. Examples of other
15 suitable fragments include the fiber having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 15. Still other suitable fragments include internal fragments. Further, the fiber protein may be modified using a variety of techniques known to those of skill in the art.

20 The invention further encompasses unique fragments of the proteins of the invention which are at least 8 amino acids in length. However, fragments of other desired lengths can be readily utilized. In addition, the invention encompasses such modifications as may be introduced to enhance yield and/or expression of an SA18 gene product, *e.g.*, construction of a fusion molecule in which all or a fragment
25 of the SA18 gene product is fused (either directly or via a linker) with a fusion partner to enhance. Other suitable modifications include, without limitation, truncation of a coding region (*e.g.*, a protein or enzyme) to eliminate a pre- or pro-protein ordinarily cleaved and to provide the mature protein or enzyme and/or mutation of a coding region to provide a secretable gene product. Still other
30 modifications will be readily apparent to one of skill in the art. The invention further encompasses proteins having at least about 95% to 99% identity to the SA18 proteins provided herein.

As described herein, vectors of the invention containing the adenoviral capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other Ad serotype based vectors, as well as other viral vectors. The rAd vectors of the invention are
5 particularly advantageous in readministration for repeat gene therapy or for boosting immune response (vaccine titers). Examples of such regimens are provided herein.

Under certain circumstances, it may be desirable to use one or more of the SA18 gene products (*e.g.*, a capsid protein or a fragment thereof) to generate an antibody. The term "an antibody," as used herein, refers to an immunoglobulin
10 molecule which is able to specifically bind to an epitope. Thus, the antibodies of the invention bind, preferably specifically and without cross-reactivity, to a SA18 epitope. The antibodies in the present invention exist in a variety of forms including, for example, high affinity polyclonal antibodies, monoclonal antibodies, synthetic antibodies, chimeric antibodies, recombinant antibodies and humanized antibodies.
15 Such antibodies originate from immunoglobulin classes IgG, IgM, IgA, IgD and IgE.

Such antibodies may be generated using any of a number of methods know in the art. Suitable antibodies may be generated by well-known conventional techniques, *e.g.* Kohler and Milstein and the many known modifications thereof. Similarly desirable high titer antibodies are generated by applying known
20 recombinant techniques to the monoclonal or polyclonal antibodies developed to these antigens [see, *e.g.*, PCT Patent Application No. PCT/GB85/00392; British Patent Application Publication No. GB2188638A; Amit *et al.*, 1986 *Science*, 233:747-753; Queen *et al.*, 1989 *Proc. Nat'l. Acad. Sci. USA*, 86:10029-10033; PCT Patent Application No. PCT/WO9007861; and Riechmann *et al.*, *Nature*, 332:323-
25 327 (1988); Huse *et al.*, 1988a *Science*, 246:1275-1281]. Alternatively, antibodies can be produced by manipulating the complementarity determining regions of animal or human antibodies to the antigen of this invention. See, *e.g.*, E. Mark and Padlin, "Humanization of Monoclonal Antibodies", Chapter 4, The Handbook of Experimental Pharmacology, Vol. 113, The Pharmacology of Monoclonal Antibodies,
30 Springer-Verlag (June, 1994); Harlow *et al.*, 1999, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow *et al.*, 1989, Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Bird *et al.*, 1988, *Science* 242:423-426.

Further provided by the present invention are anti-idiotypic antibodies (Ab2) and anti-
anti-idiotypic antibodies (Ab3). See, *e.g.*, M. Wettendorff et al., "Modulation of anti-
tumor immunity by anti-idiotypic antibodies." In *Idiotypic Network and Diseases*, ed.
by J. Cerny and J. Hiernaux, 1990 *J. Am. Soc. Microbiol.*, Washington DC: pp. 203-
5 229]. These anti-idiotypic and anti-anti-idiotypic antibodies are produced using
techniques well known to those of skill in the art. These antibodies may be used for a
variety of purposes, including diagnostic and clinical methods and kits.

Under certain circumstances, it may be desirable to introduce a
detectable label or a tag onto a SA18 gene product, antibody or other construct of the
10 invention. As used herein, a detectable label is a molecule which is capable, alone or
upon interaction with another molecule, of providing a detectable signal. Most
desirably, the label is detectable visually, *e.g.* by fluorescence, for ready use in
immunohistochemical analyses or immunofluorescent microscopy. For example,
suitable labels include fluorescein isothiocyanate (FITC), phycoerythrin (PE),
15 allophycocyanin (APC), coriphosphine-O (CPO) or tandem dyes, PE-cyanin-5 (PC5),
and PE-Texas Red (ECD). All of these fluorescent dyes are commercially available,
and their uses known to the art. Other useful labels include a colloidal gold label.
Still other useful labels include radioactive compounds or elements. Additionally,
labels include a variety of enzyme systems that operate to reveal a colorimetric signal
20 in an assay, *e.g.*, glucose oxidase (which uses glucose as a substrate) releases
peroxide as a product which in the presence of peroxidase and a hydrogen donor such
as tetramethyl benzidine (TMB) produces an oxidized TMB that is seen as a blue
color. Other examples include horseradish peroxidase (HRP) or alkaline phosphatase
(AP), and hexokinase in conjunction with glucose-6-phosphate dehydrogenase which
25 reacts with ATP, glucose, and NAD⁺ to yield, among other products, NADH that is
detected as increased absorbance at 340 nm wavelength.

Other label systems that are utilized in the methods of this invention
are detectable by other means, *e.g.*, colored latex microparticles [Bangs Laboratories,
Indiana] in which a dye is embedded are used in place of enzymes to form conjugates
30 with the target sequences provide a visual signal indicative of the presence of the
resulting complex in applicable assays.

Methods for coupling or associating the label with a desired molecule
are similarly conventional and known to those of skill in the art. Known methods of

label attachment are described [see, for example, Handbook of Fluorescent probes and Research Chemicals, 6th Ed., R. P. M. Haugland, Molecular Probes, Inc., Eugene, OR, 1996; Pierce Catalog and Handbook, Life Science and Analytical Research Products, Pierce Chemical Company, Rockford, IL, 1994/1995]. Thus,
5 selection of the label and coupling methods do not limit this invention.

The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. See, *e.g.*, Sambrook et al, Molecular Cloning: A
10 Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, **85**:2149 (1962); Stewart and Young, Solid Phase Peptide Synthesis (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of
15 skill in the art and are not a limitation of the present invention.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules. For example, in one embodiment, the simian Ad capsid
20 proteins and other simian adenovirus proteins described herein are used for non-viral, protein-based delivery of genes, proteins, and other desirable diagnostic, therapeutic and immunogenic molecules. In one such embodiment, a protein of the invention is linked, directly or indirectly, to a molecule for targeting to cells with a receptor for adenoviruses. Preferably, a capsid protein such as a hexon, penton, fiber or a fragment
25 thereof having a ligand for a cell surface receptor is selected for such targeting. Suitable molecules for delivery are selected from among the therapeutic molecules described herein and their gene products. A variety of linkers including, lipids, polyLys, and the like may be utilized as linkers. For example, the simian penton protein may be readily utilized for such a purpose by production of a fusion protein
30 using the simian penton sequences in a manner analogous to that described in Medina-Kauwe LK, et al, *Gene Ther.* **2001** May; **8**(10):795-803 and Medina-Kauwe LK, et al, *Gene Ther.* **2001** Dec; **8**(23): 1753-1761. Alternatively, the amino acid sequences of simian Ad protein IX may be utilized for targeting vectors to a cell

surface receptor, as described in US Patent Appln 20010047081. Suitable ligands include a CD40 antigen, an RGD-containing or polylysine-containing sequence, and the like. Still other simian Ad proteins, including, *e.g.*, the hexon protein and/or the fiber protein, may be used for these and similar purposes.

5 Still other adenoviral proteins of the invention may be used as alone, or in combination with other adenoviral protein, for a variety of purposes which will be readily apparent to one of skill in the art. In addition, still other uses for the adenoviral proteins of the invention will be readily apparent to one of skill in the art.

The compositions of this invention include vectors that deliver a
10 heterologous molecule to cells, either for therapeutic or vaccine purposes. Such vectors, containing simian adenovirus DNA of SA18 and a minigene, can be constructed using techniques such as those described herein for the chimeric adenoviruses and such techniques as are known in the art. Alternatively, SA19 may be a source for sequences of the chimeric adenoviruses are described herein.

15 The following examples illustrate construction and use of several chimeric viruses, including Pan5/C1, hu5/Pan7 and hu5/SV25, and Pan6/Pan7. However, these chimera are illustrative only and are not intended to limit the invention to those illustrated embodiments.

20

Example 1 - Construction of Pan5/C1 Chimeric Simian Viruses

Five different adenoviruses initially isolated from the chimpanzee, AdC68 [US Patent 6,083,716], AdPan5, AdPan7, AdPan6 and AdC1 [US Patent 6,083,716] have been sequenced. See, International Application No. PCT/US02/33645, filed
25 November 2002 for the sequences of Pan5 [SEQ ID NO:1], Pan7 [SEQ ID NO:3], and Pan6 [SEQ ID NO:2]. This application also provides sequences for SV1, SV25 and SV39 [SEQ ID No. 4, 5, 6, respectively]. Sequence comparison of the capsid protein sequences predicted that AdC1 clearly belonged to a different serological subgroup than the other four chimpanzee derived adenoviruses.

30 However, attempts to cultivate AdC1 in HEK293 cells revealed it to be fastidious in its growth characteristics (data not shown) and therefore possibly unsuitable for use as a vector using the currently available E1 complementing cell lines. However, because of the obvious sequence dissimilarity of AdC1 capsid

protein sequence from the other chimpanzee derived adenoviruses (as well as the huAd5), chimeric adenovirus vectors were generated with the capsid characteristics of AdC1. In view of the above-mentioned drawbacks associated with only making hexon changes, more extensive replacements were made in the chimera described
5 herein, *i.e.*, construction of chimeras where the replacement went beyond just the hexon, to achieve two goals. The first was to determine whether making extended replacements would allow for the rescue of viruses containing hexons of unrelated serotypes that may not otherwise be amenable to rescue. The second goal was to test whether the growth characteristics of adenovirus vectors such as AdPan5, that have
10 been found in our laboratory to be able to be grown to high titer for the purpose of manufacture, would also be present in the chimeric virus, particularly when the hexon (and other capsid proteins) are derived from a virus such as AdC1 that are difficult to grow to a high yield in cell lines such as HEK293. An added bonus of extending the replacement to include the fiber protein would be to further increase the antigenic
15 dissimilarity to beyond that afforded by a hexon change alone.

As an alternative to obtaining purified virus as source for adenoviral DNA to sequence, we have resorted to cloning restriction fragments of viral DNA obtained from infected cells ("Hirt prep"). The first adenovirus we have sequenced in this way is Simian Adenovirus. EcoRI digestion of the Simian Adenovirus yielded 7
20 fragments. Shotgun cloning yielded clones of the 5 internal fragments, which were cloned and sequenced. Completion of the sequencing was carried out by walking towards each of the ends of the genome. The map of the genome is shown in Figure 1.

A. *Construction of Two Pan5/C1 Chimeric Plasmids*

25 The overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme *Swa*I, digest the plasmid DNA with *Swa*I to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids
30 were constructed, p5C1short and p5C1long.

The plasmid p5C1short harbors an E1 deleted Pan5 virus where an internal 15226 bp segment (18332 – 33557) has been replaced by a functionally analogous 14127 bp (18531 – 32657) from AdC1. This results in the replacement of

the Pan5 proteins hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from AdC1. The ClaI site at the left end of the AdC1 fragment is at the beginning of the hexon gene and the resulting protein is identical to the C1 hexon. The EcoRI site which constitutes the right end of the AdC1 fragment is within the E4 orf 7 part of the AdC1. The right end was ligated to a PCR generated right end fragment from AdPan5 such that the regenerated orf 7-translation product is chimeric between AdPan5 and AdC1.

The plasmid p5C1long harbors an E1 deleted Pan5 virus where an internal 25603 bp segment (7955 – 33557) has been replaced by a functionally analogous 24712 bp (7946 – 32657) from AdC1. This results in the replacement of the AdPan5 pre-terminal protein, 52/55 kD protein, penton base protein, protein VII, Mu, and protein VI with those from AdC1 in addition to those replaced in p5C1short. The AscI site at the left end of the AdC1 fragment is at the beginning of the DNA polymerase gene and results in a chimeric protein where the first 165 amino acids of the AdPan5 DNA polymerase has been replaced by a 167 amino acid segment from AdC1 DNA polymerase. In this N-terminal region, the homology between the AdPan5 and AdC1 DNA polymerase proteins is 81% (72% identity).

The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction.

The plasmid pDVP5Mlu was made as follows. A synthetic DNA fragment harboring recognition sites for the restriction enzymes SmaI, MluI, EcoRI and EcoRV respectively was ligated into pBR322 digested with EcoRI and NdeI so as to retain the origin of replication and the beta-lactamase gene. The left end of Pan5 extending to the MluI site (15135 bp) was cloned into this plasmid between the SmaI and MluI sites. The E1 gene was functionally deleted and replaced by a DNA fragment harboring recognition sites for the extremely rare cutter restriction enzyme sites I-CeuI and PI-SceI. The 2904 base pairs of the right end of Pan-5 was PCR amplified using the primers P5L [GCG CAC GCG TCT CTA TCG ATG AAT TCC ATT GGT GAT GGA CAT GC, SEQ ID NO:7] and P5ITR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase (Roche). The PCR product was cut with MluI and SmaI, and cloned between MluI and EcoRV of pDVP5Mlu to yield pPan5Mlu+RE. A 3193 bp fragment extending

from the MluI site (15135) to the ClaI (18328) site of Pan5 was then inserted between the same sites of pPan5Mlu+RE to yield pPan5Cla+RE. The 3671 bp ClaI (18531) to EcoRI (22202) fragment of the adenovirus C1 was cloned into pPan5Cla+RE between ClaI (16111) and EcoRI (16116) to yield pPan5C1delRI. The 10452 bp internal
5 EcoRI fragment of the adenovirus C1 (22202 – 32653) was cloned into the EcoRI site of pPan5C1delRI to yield p5C1short. To construct p5C1long, the AdC1 replacement was further extended by replacing the AscI – ClaI 10379 bp fragment of AdPan5 in p5C1short with the AdC1 AscI – ClaI 10591 bp fragment. Finally a green fluorescent protein (GFP) expression cassette was inserted into both p5C1short and p5C1long
10 between the I-CeuI and PI-SceI sites to yield p5C1shortGFP and p5C1longGFP respectively.

B. *Rescue of chimeric Pan5/C1 recombinant vector adenoviruses*

The plasmids p5C1shortGFP and p5C1longGFP were digested with the restriction enzyme SwaI and transfected into HEK 293 cells. A typical adenovirus
15 induced cytopathic effect was observed. The rescue of recombinant chimeric adenovirus from the p5C1longGFP transfection was confirmed by collecting the supernatant from the transfection and re-infecting fresh cells which were found to be transduced as determined by GFP expression. Viral DNA prepared from the chimeric recombinant virus was digested with several restriction enzymes and found to have
20 the expected pattern on electrophoresis (data not shown).

The chimeric adenoviral construct with the shorter replacement p5C1short encodes the C1 proteins hexon and fiber as well as the intervening open reading frames for endoprotease, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, and protein VIII. (The E3 region is also included within this region
25 but is unlikely to impact on the viability of the chimeric virus). When the replacement was extended to include the additional AdC1 proteins pTP (pre-terminal protein), 52/55 kDa protein, penton base, protein VII, Mu, and protein VI, there was no difficulty in rescuing viable chimeric virus. In this experiment, the chimeric adenovirus construction strategy utilized the presence of AscI and ClaI restriction
30 enzyme sites present on the genes for DNA polymerase and hexon respectively on both AdPan5 and AdC1.

The reasons for the relatively higher yield of the chimeric virus compared to the wild-type AdC1 virus are not clear. In the growth of the 5C1

chimeric virus in 293 cells, the adenoviral early region gene products of E1 and E4 are derived from Ad5 and AdPan5 respectively. The E1 and E4 gene products bind, regulate and de-repress several cellular transcription complexes and coordinate their activity towards viral multiplication. Thus it is possible that the E1 gene products
5 supplied in *trans* from the 293 cells and the E4 gene products from AdPan5 are more optimal in the human 293 cell background than are the equivalent AdC1 gene products. This may also apply to the activity of the major late promoter whose activity is responsible for the transcription of the capsid protein genes. In the chimeric virus, the major late promoter, and the protein IVa2 which transactivates it,
10 are derived from AdPan5. However the E2 gene products required for adenoviral DNA replication pTP and single-stranded DNA – binding protein are derived from AdC1. The adenoviral DNA polymerase, which complexes with pTP, is chimeric in Ad5C1 but mostly AdPan5 derived.

15 Example 2 - Construction of Ad5 Chimeric Simian Viruses

Plasmids have been constructed where the structural proteins derive from the chimpanzee adenovirus Pan 7 and the flanking sequences are derived from human Ad5 (the commonly used vector strain). The Adhu5-Pan7 chimeric adenovirus has been rescued, demonstrating that the chimeric virus construction
20 method used to derive the chimeric virus is broadly applicable.

A. Construction of the Ad5 – Pan 7 chimeric adenovirus

A plasmid was constructed which harbors the complete (E1 deleted) chimeric genome in order to establish that the chimeric adenovirus is viable, and then transfected the plasmid into the E1 complementing cell line HEK 293. It
25 was found that the recombinant virus could be rescued. The chimeric adenovirus genome that was constructed is composed of a left end segment derived from Ad5 that contributes the ITR, the E1 deletion region containing the transgene expression cassette, the pIX and IVa2 genes and 954 C-terminal amino acids of the polymerase gene (which is transcribed in the right to left direction from the bottom strand). Ad5
30 also contributes the right end of the chimeric genome containing the E4 genes and the right ITR. All the other genes present in the central part of the chimeric construct are derived from the chimpanzee adenovirus Pan 7 including the N-terminal 235 amino acids of a chimeric DNA polymerase.

In order to construct the plasmid which harbors the complete (E1 deleted) chimeric genome, the starting plasmid was pBRAd5lere which is comprised of three parts; the bacterial origin of replication and ampicillin resistance gene derived from the plasmid pBR322, the left end of an Ad5 derived E1 deleted vector
 5 extending from the left ITR to the StuI site located at base pair number 5782 of the wild-type Ad5 genome (the E1 deletion extends from base pair 342 to 3533 of the wild-type Ad5 genome), and the right end of Ad5 extending from the StuI site at base pair number 31954 of the wild-type Ad5 genome to the right end of the right ITR. The PacI sites located adjacent to the two ITRs are used to release the Ad5 genome
 10 from the bacterial plasmid backbone. The fragment containing I-CeuI and the PI-SceI sites which is located in place of the E1 deletion is used to insert transgene cassettes.

A synthetic DNA oligomer was inserted at the StuI site containing sites for AscI, XbaI and EcoRI, which allowed the creation of the plasmid pAd5endsAscRI where using PCR, the Ad5 polymerase gene was extended to base
 15 pair #8068 of the wild-type Ad5 genome and incorporating a newly created AscI site at this location by silent mutagenesis of the polymerase gene (translated from the bottom strand) as depicted below.

Original sequence

20 GCG ACG GGC CGA [SEQ ID NO:16]
 CGC TGC CCG GCT
 Arg Arg Ala Ser [SEQ ID NO:17]

Mutated sequence (The AscI recognition site is underlined)

25 GCG GCG CGC CGA [SEQ ID NO:18]
 CGC TGC CCG GCT
 Arg Arg Ala Ser [SEQ ID NO: 17]

The Pan 7 fiber containing region was amplified by PCR (mutating the fiber stop codon from TGA to TAA to provide a polyadenylation signal similar to that in Ad5) and inserted into the EcoRI site to yield pAd5endsP7fib. Several cloning steps led to the construction of pH5C7H5 where the complete chimeric adenoviral
 30 genome has been assembled. A transgene cassette expressing GFP (green fluorescent protein) was inserted between the I-CeuI and PI-SceI sites of pH5C7H5. The final construct was digested with PacI to separate the adenoviral genome from the plasmid

backbone and transfected into HEK 293 cells. The cell lysate was harvested 2 weeks later, and the chimeric adenovirus was amplified and purified by standard methods.

5 B. Construction of the Ad5 – Simian virus 25 (SV-25) chimeric adenovirus

[N.B. Simian virus 25 (ATCC catalog number VR-201) is distinct from the chimpanzee adenovirus Simian adenovirus 25 ATCC catalog number VR-594]

10 The construction of the Ad5 based chimeric adenovirus where the left and right end segments are derived from Ad5 and the central portion was derived from the monkey adenovirus SV-25 was carried out in a manner completely analogous to that described above for the chimeric adenovirus described above that is chimeric between Ad5 and the chimpanzee adenovirus Pan 7. Thus, the chimeric adenovirus genome that was constructed is composed of a left end segment derived
15 from Ad5 that contributes the ITR, the E1 deletion region containing the transgene expression cassette, the pIX and IVa2 genes and 956 C-terminal amino acids of the polymerase gene. Ad5 also contributes the right end of the chimeric genome containing the E4 genes and the right ITR. [Additionally, the left end of the Ad5 genome was extended beyond that present in pH5C7H5 so that 454 base pairs of the
20 Ad5 left end was present. Although not absolutely essential, this was done in order to improve packaging efficiency.] All the genes present in the central part of the chimeric construct are derived from the monkey adenovirus SV-25 including the N-terminal 230 amino acids of a chimeric DNA polymerase. The starting plasmid for the construction of the chimeric genome was pAd5endsAscRI which contains both
25 the left and right ends of Ad5 as well as the created (by silent mutation) AscI site in the polymerase gene where Ad5-SV25 chimeric fusion was made (as was done for the Ad5 – Pan 7 chimeric adenovirus). In the final construct pH5S25H5, the SV25 genome segment has been incorporated by sequential cloning steps, including creation of an AscI site at the ligation junction within the polymerase coding
30 sequence. A transgene cassette expressing GFP (green fluorescent protein) was inserted between the I-CeuI and PI-SceI sites of pH5S25H5. The final construct was digested with PacI to separate the adenoviral genome from the plasmid backbone and

transfected into HEK 293 cells. The cell lysate was harvested 2 weeks later, and the chimeric adenovirus was amplified and purified by standard methods.

Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus. The portion of the genome replaced by DNA from Pan7 is indicated.

5

Example 3 – Pan5 - C1 chimeric vector of invention as a delivery vehicle for immunogenic compositions

A Pan 5 (Simian adenovirus 22, a subgroup E adenovirus, also termed C5) - C1 (Simian adenovirus 21, a subgroup B adenovirus) chimeric expressing the Ebola virus (Zaire) glycoprotein (C5C1C5-CMVGP) was constructed as a model antigen in order to test the efficacy of the vector C5C1C5-CMVGP as a vaccine; this vector has been compared it to the Adhu5 based vector (H5-CMVGP). Compared to H5-CMVGP, the C5C1-CMVGP vector yielded only a slightly decreased level of GP expression in transduced A549 cells.

Thereafter, GP-specific T cell and B cell responses elicited in B10BR mice vaccinated intramuscularly with either 5×10^{10} H5-CMVGP or C5C1-CMVGP vectors were compared.

The C5C1C5-CMVGP vector appeared to induce lower frequencies of gamma interferon producing CD8⁺ T cells with kinetics slower than the H5-CMVGP vector as determined by intracellular cytokine staining using a H-2k restricted GP-specific peptide as stimulant. The total IgG response to GP, measured by ELISA, was equivalent in serum from mice vaccinated with the C5C1C5-CMVGP or the H5-CMVGP vectors. However, the C5C1C5-CMVGP vector induced a more potent Th1 type response while the H5-CMVGP vector stimulated a more balanced Th1/Th2 type response. In a survival study, mice were vaccinated as above and challenged 28 days later with 200 LD/50 mouse-adapted Ebola Zaire virus. 100% survival was seen for both groups.

Example 4 – Generation of Chimeric Pan6/Pan7 Vectors

A panel of GFP expressing vectors were generated. This panel includes vectors that are chimeric between Pan 6 and Pan 7 where (a) the hexon protein of Pan 7 was replaced by that of Pan 6 (termed C767), (b), the fiber protein of Pan7 was

replaced by that of Pan 6 (termed C776), (c) both the hexon and fiber proteins of the Pan 7 vector have been replaced by those from Pan 6 (termed C766).

The chimeric virus termed C767 was constructed essentially as described above for the C5C1C5 virus in Example 1. However, due to substantial homology
5 between the Pan6 and Pan7 sequences 5' to the hexon sequence, it was not necessary to substitute the 5' end of the genome between the penton and the pol gene.

The chimeric vector C767 was compared to the C776, C766, the parent C6, and the parent C7, each expressing GFP.

Balb/C mice (25 per group) were immunized intramuscularly with either
10 Pan 6 or Pan 7 (10^{10} particles). Re-administration (10^{11} particles i.v., by tail vein injection) was attempted 3 weeks later using each of the five GFP expressing vectors (C6-GFP, C7-GFP, and the three chimeric vectors). Three days later the level of liver transduction was estimated qualitatively by examining liver sections for the presence of GFP expression and quantitatively by estimating copies of GFP DNA by Taqman
15 analysis. Administration of either one of the two chimpanzee adenovirus vectors does not affect the transduction efficiency of the other vector, while re-administration of the same vector is severely compromised. The data showed that antibodies to both hexon and fiber are important in preventing re-administration of adenoviral vectors.

20

All publications cited in this specification are incorporated herein by reference, as are the priority documents, US Patent Application 60/575,429, filed March 28, 2004; US Patent Application No. 60/566,212, filed April 28, 2004, and US Patent Application No. 10/465,302, filed June 20, 2003. While the invention has been
25 described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

What is claimed is:

1. A method of efficiently culturing a chimeric adenovirus in a selected host cell, said chimeric adenovirus being from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the steps of:

(a) generating a chimeric adenovirus comprising:

(i) adenovirus sequences of the left terminal end and right terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the 5' inverted terminal repeat (ITRs), and said right end region comprising the 3' inverted terminal repeat (ITRs); and

(ii) the internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the late genes encoding the penton, hexon, and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', a left terminal region of the first adenovirus, the internal region of the parental adenovirus, and the right terminal region of the first adenovirus; and

b) culturing said chimeric adenovirus in the presence of functional adenovirus E1a, E1b, and E4 ORF6 genes from the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus, and further in the presence of necessary adenoviral structural genes from the left end of the adenovirus.

2. The method according to claim 1, wherein the internal region of the parental adenovirus further comprises one or more functional adenovirus genes selected from the group consisting of Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and protein VI.

3. The method according to claim 1, wherein the polymerase, terminal protein and 52/55 kDa protein functions are provided in *trans*.

4. The method according to claim 1, wherein the first adenovirus further comprises the polymerase, terminal protein and 52/55 kDa protein functions.
5. The method according to claim 1, wherein the chimeric adenovirus comprises the adenoviral late genes 1, 2, 3, 4, and 5 of the parental adenovirus.
6. The method according to claim 1, wherein the selected host cell stably contains one or more of the adenovirus E1a, E1b or E4 ORF6 functions.
7. The method according to claim 1, wherein the chimeric adenovirus comprises one or more of the adenovirus E1a, E1b or E4 ORF6 of the first adenovirus.
8. The method according to claim 1, wherein the first adenovirus is of human origin.
9. The method according to claim 1, wherein the first adenovirus is of simian origin.
10. The method according to claim 1, further comprising the step of isolating the chimeric adenovirus.
11. A method for generating a chimeric adenovirus for growth in a selected host cell, said chimeric adenovirus being derived from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the step of generating a chimeric adenovirus comprising:
 - 5' and 3' terminal regions of a first adenovirus which grows in a selected host cell type, said 5' terminal regions comprising the 5' inverted terminal repeat (ITRs) and necessary E1 gene functions, and said 3' terminal regions comprising inverted terminal repeat (ITRs) and necessary E4 gene functions; and
 - internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the hexon, penton base and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', the 5' terminal region of the first adenovirus, the internal region of the parental adenovirus, and the 3' terminal regions of the first adenovirus.

12. A chimeric adenovirus produced according to the method of claim 1.

13. A chimeric adenovirus comprising a hexon protein of a selected adenovirus serotype which is incapable of efficient growth in a selected host cell, said modified adenovirus comprising:

(a) adenovirus sequences of the left terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the E1a, E1b and 5' inverted terminal repeat (ITRs);

(b) adenovirus sequences of the internal region of the selected adenovirus serotype which is incapable of efficient growth in the selected host cell, said internal region comprising the genes encoding the penton, hexon and fiber of the selected adenovirus;

(c) adenovirus sequences of the right terminal end of the first adenovirus, said right end region comprising the necessary E4 gene functions and the 3' inverted terminal repeat (ITRs),

wherein the resulting chimeric adenovirus comprises adenoviral structural and regulatory proteins necessary for infection and replication.

14. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus further comprises the IIIa, 52/55kDa and terminal protein (pTP) of the selected adenovirus serotype.

15. The chimeric adenovirus according to claim 13, wherein chimeric adenovirus comprises the polymerase of the first adenovirus.

16. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus expresses a functional chimeric protein formed from the first adenovirus and the selected adenovirus, said chimeric protein is selected from the group consisting of polymerase, terminal protein, 52/55 kDa protein, and IIIa.

17. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus comprises the terminal protein, 52/55 kDa, and/or IIIa of the selected adenovirus.

18. A host cell comprising a chimeric adenovirus according to claim 12.

19. The host cell according to claim 18, wherein said host cell is a human cell.

20. An isolated simian adenovirus nucleic acid sequence selected from the group consisting of:

(a) SA18 having the sequence of nucleic acids 1 to 31967 of SEQ ID NO:12 and

(b) a nucleic acid sequence complementary to the sequence of any of (a) to (f).

21. An isolated simian adenovirus serotype nucleic acid sequence selected from one or more of the group consisting of:

(a) 5' inverted terminal repeat (ITR) sequences;

(b) the adenovirus E1a region, or a fragment thereof selected from among the 13S, 12S and 9S regions;

(c) the adenovirus E1b region, or a fragment thereof selected from among the group consisting of the small T, large T, IX, and IVa2 regions;

(d) the E2b region;

(e) the L1 region, or a fragment thereof selected from among the group consisting of the 28.1 kD protein, polymerase, agnoprotein, 52/55 kD protein, and IIIa protein;

(f) the L2 region, or a fragment thereof selected from the group consisting of the penton, VII, VI, and Mu proteins;

(g) the L3 region, or a fragment thereof selected from the group consisting of the VI, hexon, or endoprotease;

(h) the 2a protein;

(i) the L4 region, or a fragment thereof selected from the group consisting of the 100 kD protein, the 33 kD homolog, and VIII;

(j) the E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;

(k) the L5 region, or a fragment thereof selected from a fiber protein;

(l) the E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1; and

(m) the 3' ITR, of any of SA18 SEQ ID NO:12, or a sequence complementary to any of (a) to (m).

22. A simian adenovirus protein encoded by the nucleic acid sequence according to claim 21.

23. A composition comprising a simian adenovirus capsid protein according to claim 22 linked to a heterologous molecule for delivery to a selected host cell.

24. A method for targeting a cell having an adenoviral receptor comprising delivering to a subject a composition according to claim 23.

25. A nucleic acid molecule comprising a heterologous simian adenoviral sequence according to claim 21.

26. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence encodes an adenoviral gene product and is operatively linked to regulatory control sequences which direct expression of the adenoviral gene product in a host cells.

27. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence comprises the E1a region of SA18 SEQ ID NO:12.

28. A pharmaceutical composition comprising the nucleic acid molecule according to claim 27 and a physiologically compatible carrier.
29. A recombinant adenovirus having a capsid comprising a protein selected from the group consisting of:
- (a) a hexon protein of SA18, SEQ ID NO 13, or a unique fragment thereof;
 - (b) a penton protein of SA18, SEQ ID NO: 14, or a unique fragment thereof;
 - (c) a fiber protein of SA18, SEQ ID NO: 15, or a unique fragment thereof.
30. The recombinant adenovirus according to claim 29, wherein the capsid is of an artificial serotype.
31. The recombinant adenovirus according to claim 29, wherein said virus further comprises a heterologous gene operatively linked to sequences which direct expression of said gene in a host cell.
32. The recombinant adenovirus according to claim 29, further comprising 5' and 3' adenovirus cis-elements necessary for replication and encapsidation.
33. The recombinant adenovirus according to claim 29, wherein said vector lacks all or a part of the E1 gene.
34. A host cell comprising a heterologous nucleic acid molecule comprising the nucleic acid sequence according to claim 21.
35. The host cell according to claim 34, wherein said host cell is stably transformed with the nucleic acid molecule.

36. The host cell according to claim 34, wherein said host cell expresses one or more adenoviral gene products from said nucleic acid molecule, said adenoviral gene products selected from the group consisting of E1a, E1b, E2a, and E4 ORF6.

37. The host cell according to claim 34, wherein said host cell is stably transformed with a nucleic acid molecule comprising the simian adenovirus inverted terminal repeats.

38. A composition comprising a recombinant virus according to claim 29 in a pharmaceutically acceptable carrier.

39. A method for delivering a heterologous gene to a mammalian cell comprising introducing into said cell an effective amount of the recombinant virus according to claim 29.

40. A method for repeat administration of a heterologous gene to a mammal comprising the steps of:

- (a) introducing into said mammal a first vector which comprises the heterologous gene and
- (b) introducing into said mammal a second vector which comprises the heterologous gene;
wherein at least the first virus or the second vector is a virus according to claim 29 and wherein the first and second recombinant vector are different.

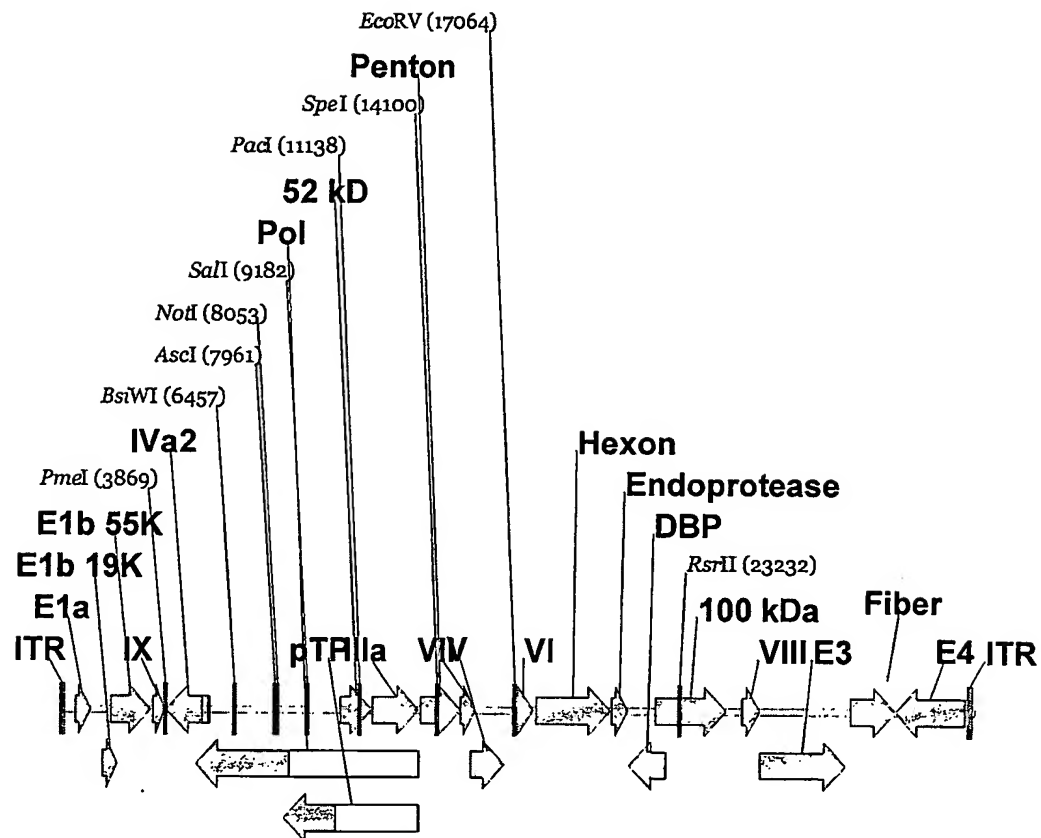
41. A method for producing a selected gene product comprising infecting a mammalian cell with the recombinant virus according to claim 29, culturing said cell under suitable conditions and recovering from said cell culture the expressed gene product.

42. A method for eliciting an immune response in a mammalian host against an infective agent comprising administering to said host an effective amount of the

recombinant adenovirus of claim 29, wherein said heterologous gene encodes an antigen of the infective agent.

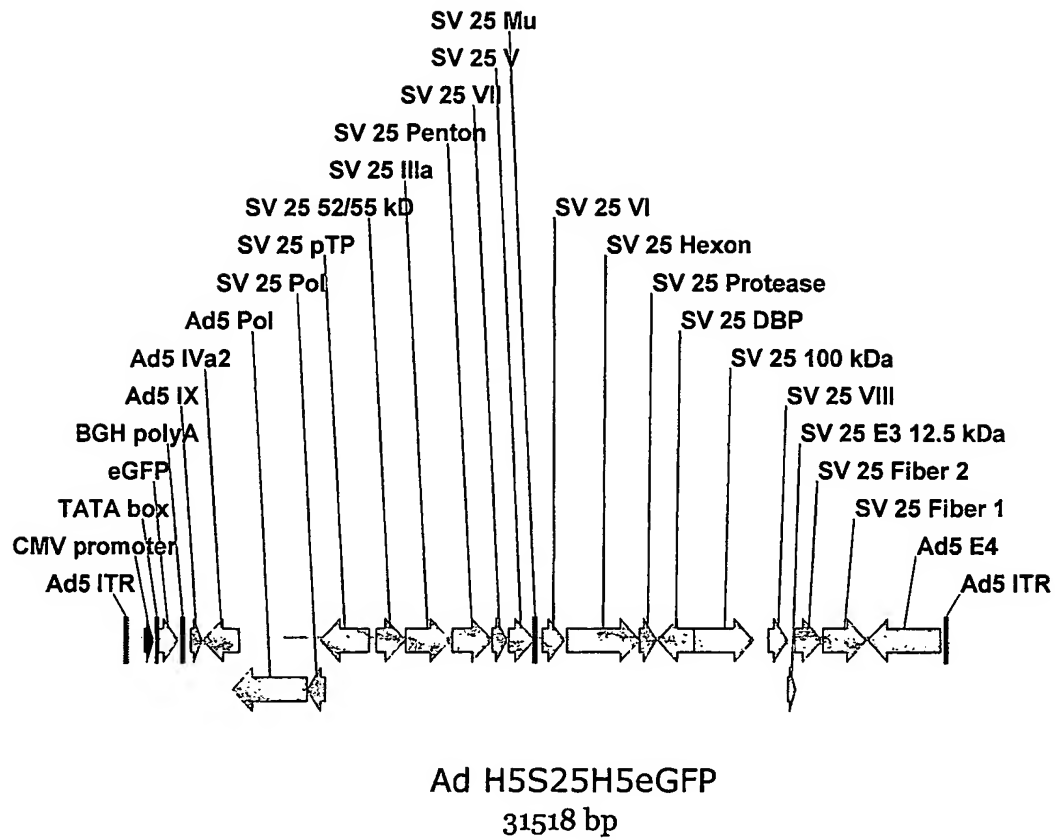
43. The method according to claim 42, comprising the step of priming the host with a DNA vaccine comprising the heterologous gene prior to administering the recombinant adenovirus.

Fig. 1



Simian Adenovirus
34302 bp

Fig. 2



SEQUENCE LISTING

<110> The Trustees of the University of Pennsylvania
 Roy, Soumitra
 Wilson, James M.

<120> Methods of Generating Chimeric Adenoviruses and Uses For Such
 Chimeric Adenoviruses

<130> UPN-P3067PCT

<150> US 10/465,302
 <151> 2003-06-20

<150> US 60/566,212
 <151> 2004-04-28

<150> US 60/575,429
 <151> 2004-05-28

<160> 18

<170> PatentIn version 3.2

<210> 1
 <211> 36462
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan5

<400> 1
 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaag aggtatttga 60
 atttggggat gcggggcggg gattggctgc gggagcggcg accgttaggg gcggggcggg 120
 tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
 ggaaatgagg tgtttctggg cggtatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggt gtcagctgat cgccagggtg 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
 gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg 660
 gtgacgaccc tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgtatg 720
 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
 gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
 cctctctcca taccgccaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggcttttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
 acagtaagtg tgattaactt tagctgggga ggcagagggt gactgggtgc tgactggttt 1200

atztatgtat	atgtttttta	tgtgtaggtc	ccgtctctga	cgtagatgag	acccccacta	1260
cagagtgc	ttcatcacc	ccagaaattg	gcgaggaacc	gcccgaagat	attattcata	1320
gaccagttgc	agtgagagtc	accgggcgta	gagcagctgt	ggagagtttg	gatgacttgc	1380
tacaggggtg	ggatgaacct	ttggacttgt	gtacccggaa	acgccccagg	cactaagtgc	1440
cacacatgtg	tgtttactta	aggtgatgtc	agtatttata	gggtgtggag	tgcaataaaa	1500
tccgtgttga	ctttaagtgc	gtggtttatg	actcaggggt	ggggactgtg	ggtatataag	1560
caggtgcaga	cctgtgtggt	cagttcagag	caggactcat	ggagatctgg	acagtcttgg	1620
aagactttca	ccagactaga	cagctgctag	agaactcatc	ggagggagtc	tcttacctgt	1680
ggagattctg	cttcgggtgg	cctctagcta	agctagtcta	tagggccaag	caggattata	1740
aggatcaatt	tgaggatatt	ttgagagagt	gtcctggtat	ttttgactct	ctcaacttgg	1800
gccatcagtc	tcactttaac	cagagtattc	tgagagccct	tgacttttct	actcctggca	1860
gaactaccgc	cgcggtagcc	ttttttgcct	ttatccttga	caaattggagt	caagaaaccc	1920
atttcagcag	ggattaccgt	ctggactgct	tagcagtagc	tttgtggaga	acatggaggt	1980
gccagcgct	gaatgcaatc	tccggctact	tgccagtaca	gccggtagac	acgctgagga	2040
tcctgagtct	ccagtcaccc	caggaacacc	aacgccgcca	gcagccgcag	caggagcagc	2100
agcaagagga	ggaccgagaa	gagaacctga	gagccggtct	ggaccctccg	gtggcggagg	2160
aggaggagta	gctgacttgt	ttcccagact	gcgccgggtg	ctgactaggt	cttcagtggt	2220
acgggagagg	gggattaagc	gggagaggca	tgaggagact	agccacagaa	ctgaactgac	2280
tgctcagtctg	atgagtcgca	ggcgcccaga	atcggtgtgg	tggcatgagg	tgagtcgca	2340
ggggatagat	gaggtctcag	tgatgcatga	gaaatattcc	ctagaacaag	tcaagacttg	2400
ttggttggag	cccagggatg	attgggaggt	agccatcagg	aattatgcca	agctggctct	2460
gaggccagac	aagaagtaca	agattaccaa	actgattaat	atcagaaatt	cctgctacat	2520
ttcagggaat	ggggccgagg	tggagatcag	taccagagg	agggtggcct	tcagatgctg	2580
catgatgaat	atgtacccgg	gggtggtggg	catggaggga	gtcaccttta	tgaacgcgag	2640
gttcaggggt	gatgggtata	atggggtggt	ctttatggcc	aacaccaagc	tgacagtgca	2700
cggatgctcc	ttctttggct	tcaataacat	gtgcattgag	gcctggggca	gtgtttcagt	2760
gaggggatgc	agtttttcag	ccaactggat	gggggtcgtg	ggcagaacca	agagcatggt	2820
gtcagtgaag	aatgcctgt	tcgagaggtg	ccacctgggg	gtgatgagcg	agggcgaagc	2880
caaagtcaaa	cactgcgct	ctaccgagac	gggctgcttt	gtactgatca	agggcaatgc	2940
caaagtcaag	cataatatga	tctgtggggc	ctcggatgag	cgcggtacc	agatgctgac	3000
ctgcgccggt	gggaacagcc	atatgctagc	caccgtgcat	gtggcctcgc	acccccgcaa	3060
gacatggccc	gagttcgagc	acaacgtcat	gacccgctgc	aatgtgcacc	tggggctccg	3120
ccgaggcatg	ttcatgccct	accagtgcaa	catgcaattt	gtgaaggtgc	tgctggagcc	3180
cgatgccatg	tccagagtga	gcctgacggg	ggtgtttgac	atgaatgtgg	agctgtggaa	3240

aattctgaga	tatgatgaat	ccaagaccag	gtgccggggc	tgcaatgcg	gaggcaagca	3300
cgccaggctt	cagcccgtgt	gtgtggaggt	gacggaggac	ctgcgacccg	atcatttggt	3360
gttgtcctgc	aacgggacgg	agttcggtc	cagcggggaa	gaatctgact	agagtgahta	3420
gtgtttggga	ctgggtggga	gcctgcatga	tgggcagaat	gactaaaatc	tgtgtttttc	3480
tgcgacgag	catgagcgga	agcgctcct	ttgagggagg	ggtattcagc	ccttatctga	3540
cggggcgctt	cccctcctgg	gcgggagtg	gtcagaatgt	gatgggatcc	acggtggacg	3600
gccggcccg	gcagcccgcg	aactcttcaa	ccctgacct	cgcgaccctg	agctcctcgt	3660
ccgtggacgc	agctgccgcc	gcagctgctg	cttcgcgcgc	cagcgccgtg	cgcggaatgg	3720
ccctgggcgc	cggctactac	agctctctgg	tggccaactc	gagttccacc	aataatcccg	3780
ccagcctgaa	cgaggagaag	ctgctgctgc	tgatggccca	gctcgaggcc	ctgaccacgc	3840
gcctgggcga	gctgaccacg	caggtggctc	agctgcaggc	ggagacgcgg	gccgcgggtg	3900
ccacggtgaa	aaccaaataa	aaaatgaatc	aataaataaa	cggagacggt	tgttgatttt	3960
aacacagagt	cttgaatctt	tatttgattt	ttcgcgcgcg	gtaggccctg	gaccaccggt	4020
ctcgatcatt	gagcaccgcg	tggatctttt	ccaggaccgc	gtagaggtgg	gcttgatgtg	4080
tgaggtacat	gggcatgagc	ccgtcccggg	ggtggaggta	gctccattgc	agggcctcgt	4140
gctcgggggt	ggtgttgtaa	atcaccacgt	catagcaggg	gcgcagggcg	tgggtctgca	4200
cgatgtcctt	gaggaggaga	ctgatggcca	cgggcagccc	cttgggtgtag	gtgttgacga	4260
acctgttgag	ctgggaggga	tgcatgcggg	gggagatgag	atgcatcttg	gcctggatct	4320
tgagattggc	gatgttcccg	cccagatccc	gccggggggt	catgttgctg	aggaccacca	4380
gcacggtgta	tccggtgcac	ttggggaatt	tgtcatgcaa	cttgggaagg	aaggcgtgaa	4440
agaatttgga	gacgcccttg	tgaccgccc	ggttttccat	gcactcatcc	atgatgatgg	4500
cgatggggcc	gtgggcggcg	gcttgggcaa	agacgtttcg	ggggtcggac	acatcgtagt	4560
tgtggtcctg	ggtgagctcg	tcataggcca	ttttaatgaa	tttggggcgg	agggtgcccg	4620
actgggggac	gaaggtgccc	tcgatcccgc	gggcgtagtt	gccctcgag	atctgcatct	4680
cccaggcctt	gagctcggag	ggggggatca	tgtccacctg	cggggcgatg	aaaaaacgg	4740
tttccggggc	gggggagatg	agctgggccc	aaagcaggtt	ccggagcagc	tgggacttgc	4800
cgcagccggt	ggggccgtag	atgaccccca	tgaccggctg	caggtggtag	ttgagggaga	4860
gacagctgcc	gtcctcgcgc	aggagggggg	ccacctcggt	catcatctcg	cgacatgca	4920
tgttctcgcg	cacgagttcc	gccaggaggc	gctcgcccc	aagcgagagg	agctcttgca	4980
gcgaggcgaa	gtttttcagc	ggcttgagcc	cgctggccat	gggcattttg	gagaggggtct	5040
gttgcaagag	ttccagacgc	tcccagagct	cggtgatgtg	ctctagggca	tctcgatcca	5100
gcagacctcc	tcgtttcgcg	ggttggggcg	actgcgggag	tagggcacca	ggcgatgggc	5160
gtccagcgag	gccagggtcc	ggctcttcca	ggggcgaggg	gtccgcgtca	gcgtggtctc	5220
cgtcacgggtg	aaggggtgcg	cgccgggctg	ggcgcttgcg	aggggtgcgt	tcaggctcat	5280
ccggctgggtc	gagaaccgct	cccggtcggc	gccctgcgcg	tcggccaggt	agcaattgag	5340

catgagttcg	tagttgagcg	cctcggccgc	gtggcccttg	gcgcggagct	tacctttgga	5400
agtgtgtccg	cagacgggac	agaggagggg	cttgagggcg	tagagcttgg	gggcgaggaa	5460
gacggactcg	ggggcgtagg	cgtccgcgcc	gcagctggcg	cagacggtct	cgactccac	5520
gagccaggtg	aggtctggcc	ggtcggggtc	aaaaacgagg	tttcctccgt	gctttttgat	5580
gcgtttctta	cctctggtct	ccatgagctc	gtgtccccgc	tgggtgacaa	agaggctgtc	5640
cgtgtccccg	tagaccgact	ttatggggcg	gtcctcgagc	ggggtgccgc	ggtcctcgtc	5700
gtagaggaac	ccgcccact	ccgagacgaa	ggcccgggtc	caggccagca	cgaaggaggc	5760
cacgtgggag	gggtagcggg	cgttggtccac	cagcgggtcc	accttctcca	gggtatgcaa	5820
gcacatgtcc	ccctcgcca	catccaggaa	ggtgattggc	ttgtaagtgt	aggccacgtg	5880
accgggggtc	ccggccgggg	gggtataaaa	gggggagggc	ccctgctcgt	cctcactgtc	5940
ttccggatcg	ctgtccagga	gcgccagctg	ttggggtagg	tattccctct	cgaaggcggg	6000
catgacctcg	gactcaggt	tgtcagtttc	tagaaacgag	gaggatttga	tattgacggg	6060
gccgttggag	acgcctttca	tgagcccctc	gtccatctgg	tcagaaaaga	cgatcttttt	6120
gttgtcgagc	ttggtggcga	aggagccgta	gagggcggtg	gagagcagct	tggcgatgga	6180
gcgcatggtc	tggttctttt	ccttgtcggc	gcgctccttg	gcggcgatgt	tgagctgcac	6240
gtactcgcg	gccacgcact	tccattcggg	gaagacggtg	gtgagcttgt	cgggcacgat	6300
tctgaccgc	cagccgcggg	tgtgcagggt	gatgagggtc	acgctgggtg	ccacctcgcc	6360
gcgcaggggc	tcgttgggtc	agcagaggcg	cccggccttg	cgcgagcaga	aggggggcag	6420
cgggtccagc	atgagctcgt	cgggggggtc	ggcgtccacg	gtgaagatgc	cgggcaggag	6480
ctcggggtcg	aagtagctga	tgcagggtgc	cagatcgtcc	agcgccgctt	gccagtcgcg	6540
cacggccagc	gcgcgctcgt	aggggctgag	gggcgtgccc	cagggcatgg	ggtgcgtgag	6600
cgcgagggcg	tacatgccgc	agatgtcgta	gacgtagagg	ggctcctcga	ggacgccgat	6660
gtaggtgggg	tagcagcgcc	ccccgcggat	gctggcgcg	acgtagtcgt	acagctcgtg	6720
cgagggcgcg	aggagcccgg	tgccgaggtt	ggagcgctgc	ggcttttcgg	cgcggtagac	6780
gatctggcgg	aagatggcgt	gggagttgga	ggagatggtg	ggcctctgga	agatgttgaa	6840
gtgggcgtgg	ggcagtcgca	ccgagtcctt	gatgaagtgg	gcgtaggagt	cctgcagctt	6900
ggcgacgagc	tcggcggtga	cgaggacgtc	cagggcgag	tagtcgaggg	tctcttggtg	6960
gatgtcgtac	ttgagctggc	ccttctgctt	ccacagctcg	cggttgagaa	ggaactcttc	7020
gcggtccttc	cagtactctt	cgagggggaa	cccgtcctga	tcggcacggg	aagagcccac	7080
catgtagaac	tggttgacgg	ccttgtaggc	gcagcagccc	ttctccacgg	ggagggcgta	7140
agcttgcgcg	gccttgcgca	gggaggtgtg	ggtgagggcg	aagggtgtcg	gcaccatgac	7200
cttgaggaac	tgggtgcttg	agtcgaggtc	gtcgcagccg	ccctgctccc	agagctggaa	7260
gtccgtgcgc	ttctttaggg	cgggggtggg	caaagcgaaa	gtaacatcgt	tgaagaggat	7320
cttgcccgcg	cggggcatga	agttgcgagt	gatgcggaaa	ggctggggca	cctcggcccc	7380

gttggtgatg	acctgggagg	cgaggacgat	ctcgtcgaag	ccgttgatgt	tgtgcccagc	7440
gatgtagagt	tccacgaatc	gcgggaggcc	cttgacgtgg	ggcagcttct	tgagctcgtc	7500
gtaggtgagc	tcggcggggg	cgctgaggcc	gtgctgctcg	agggccaggt	cggcgaggtg	7560
gggggttgcg	ccgaggaagg	aagtccagag	atccacggcc	agggcgggtc	gcaagcggtc	7620
ccggtactga	cggaactgct	ggcccacggc	cattttttcg	ggggtgacgc	agtagaaggt	7680
gcgggggtcg	ccgtgccagc	ggccccactt	gagctggagg	gcgaggtcgt	gggcgagctc	7740
gacgagcggc	gggtccccgg	agagtttcat	gaccagcatg	aaggggacga	gctgcttgcc	7800
gaaggacccc	atccaggtgt	aggtttccac	gtcgtaggtg	aggaagagcc	tttcgggtgcg	7860
aggatgcgag	ccgatgggga	agaactggat	ctcctgccac	cagttggagg	aatggctgtt	7920
gatgtgatgg	aagtagaaat	gccgacggcg	cgccgagcac	tcgtgcttgt	gtttatacaa	7980
gcgtccgcag	tgctcgcaac	gctgcacggg	atgcacgtgc	tgacagagct	gtacctgggt	8040
tcctttgacg	aggaatttca	gtgggcagtg	gagcgctggc	ggctgcatct	ggtgctgtac	8100
tacgtcctgg	ccatcggcgt	ggccatcgtc	tgccctcatg	gtggctcatgc	tgacgaggcc	8160
gcgcgggagg	caggtccaga	cctcggtctg	gacgggtcgg	agagcgagga	cgagggcgcg	8220
caggccggag	ctgtccaggg	tcctgagacg	ctgcggagtc	aggtcagtg	gcagcggcgg	8280
cgcgcggttg	acttgacagga	gcttttccag	ggcgcgcggg	aggccagat	ggtacttgat	8340
ctccacggcg	ccgttggtgg	cgacgtccac	ggcttgacag	gtcccgtgcc	cctggggcgc	8400
caccaccgtg	ccccgtttct	tcttggtgtc	tggcgggcgg	ggctccatgc	ttagaagcgg	8460
cggcgaggac	gcgcgcgggg	cggcaggggc	ggctcggggc	ccggaggcag	gggcggcagg	8520
ggcacgtcgg	cgccgcgcgc	gggcaggttc	tgggtactgc	cccggagaag	actggcgtga	8580
gcgacgacgc	gacggttgac	gtcctggatc	tgacgcctct	gggtgaaggc	cacgggaccc	8640
gtgagtttga	acctgaaaga	gagttcgaca	gaatcaatct	cggtatcgtt	gacggcggcc	8700
tgccgcagga	tctcttgac	gtcgcccgag	ttgtcctgg	aggcgatctc	ggtcatgaac	8760
tgctcgtatc	cctcctcctg	aaggctctcc	cgaccggcgc	gctcgacggt	ggccgcgagg	8820
tcgttgagga	tgcgcccat	gagctgcgag	aaggcgttca	tgccggcctc	gttccagacg	8880
cggctgtaga	ccacggctcc	gtcggggctg	cgcgcgcgca	tgaccacctg	ggcgaggttg	8940
agctcgacgt	ggcgcggtga	gaccgcgtag	ttgcagaggc	gctggtagag	gtagttgagc	9000
gtggtggcga	tgtgctcgg	gacgaagaag	tacatgatcc	agcggcgag	cggcatctcg	9060
ctgacgtcgc	ccagggcttc	caagcgctcc	atggcctcgt	agaagtccac	ggcgaagttg	9120
aaaaactggg	agttgcgcgc	cgagacggtc	aactcctcct	ccagaagacg	gatgagctcg	9180
gcgatggtgg	cgcgcacctc	gcgctcgaag	gccccggggg	gctcctcttc	ttccatctcc	9240
tcctcctctt	ccatctcctc	cactaacatc	tcttctactt	cctcctcagg	aggcggcggc	9300
gggggagggg	ccctgcgtcg	ccggcggcgc	acgggcagac	ggtcgatgaa	gcgctcgtatg	9360
gtctccccgc	gccggcgacg	catggtctcg	gtgacggcgc	gcccgtcctc	gcggggccgc	9420
agcgtgaaga	cgccgccgcg	catctccagg	tggccgccgg	gggggtctcc	gttgggcagg	9480

gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg 9540
agcgtctcga gatccacggg atccgaaaac cgctgaacga aggcttcgag ccagtcgcag 9600
tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg 9660
gcgatgctgc tggatgatgaa gttgaagtag gcggtcctga gacggcggat ggtggcgagg 9720
agcaccaggt ccttggggccc ggcttgctgg atgcgacagc ggtcggccat gccccaggcg 9780
tggtcctgac acctggcgag gtccttgtag tagtcctgca tgagccgctc cacgggcacc 9840
tcctcctcgc ccgcgcggcc gtgcatgcgc gtgagcccga acccgcgctg cggctggacg 9900
agcgccaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgaggggtg 9960
gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cgggtgttgat ggtgtaggag 10020
cagttggcca tgacggacca gttgacggtc tggtagggcg ggcgcacgag ctctgtggtac 10080
ttgagggcgc agtagggcgc cgtgtcgaag atgtagtcgt tgcaggtgcg cacgaggtac 10140
tggtatccga cgaggaagtg cggcggcggc tggcggtaga gcggccatcg ctcggtggcg 10200
ggggcgccgg gcgcgaggtc ctcgagcatg aggcggtggt agccgtagat gtacctggac 10260
atccaggtga tgccggcggc ggtggtggag gcgcgcggga actcgcgagc gcggttccag 10320
atgttgcgca gcggcaggaa gtagttcatg gtggccgcgg tctggcccgt gaggcgcgcg 10380
cagtcgtgga tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactccgt 10440
ggcctggagg ctaagcgaac gggttgggct gcgcgtgtac cccggttcga gtccctgctc 10500
gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc 10560
taacgaaacc tccaggatac ggaggcgggt cgttttggcc attttcgtca ggccggaaat 10620
gaaactagta agcgcggaac gcggccgtcc gcgatggctc gctgccgtag tctggagaaa 10680
gaatcgccag ggttgcggtt cggtgtgccc cggttcgagc ctacgcgctc ggcgcgggcc 10740
ggattccgcg gctaacgtgg gcgtggctgc cccgtcgttt ccaagacccc ttagccagcc 10800
gacttctcca gttacggagc gagcccctct ttttcttgtg tttttgccag atgcatcccg 10860
tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac 10920
agccggcgct tctgcccccg cccagcagc agcagccagc cactaccgcg gcggccgccg 10980
tgagcggagc cggcgttcag tatgacctgg ccttggaaga gggcgagggg ctggcgcggc 11040
tgggggcgct gtcgcccggag cggcaccgc gcgtgcagat gaaaaggagc gctcgcgagg 11100
cctacgtgcc caagcagaac ctgttcagag acaggagcgg cgaggagccc gaggagatgc 11160
gcgcctcccg cttccacgcg gggcgggagc tgccggcgcg cctggaccga aagcgggtgc 11220
tgagggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgacg 11280
tgcccgcggc caacctggtc acggcgtagc agcagaccgt gaaggaggag agcaacttcc 11340
aaaaatcctt caacaaccac gtgcgcacgc tgatcgcgcg cgaggagggtg accctgggccc 11400
tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa cccacgagc aagccgctga 11460
cggcgcagct gtttctggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc 11520

tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga 11580
gcatcgtggt gcaggagcgc gggctgccgc tgtccgagaa gctggcggcc atcaacttct 11640
cggtgctgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgccc 11700
tagacaagga ggtgaagatc gacgggtttt acatgcgcat gaccctgaaa gtgctgaccc 11760
tgagcgacga tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgccagcc 11820
gccggcgca gctgagcgac caggagctga tgcacagcct gcagcgggccc ctgaccgggg 11880
ccgggaccga gggggagagc tactttgaca tgggcgcgga cctgcgctgg cagcctagcc 11940
gccgggcctt ggaagctgcc ggcggttccc cctacgtgga ggaggtggac gatgaggagg 12000
aggagggcga gtacctggaa gactgatggc gcgaccgtat ttttgctaga tgcagcaaca 12060
gccaccgccg cctcctgatc ccgcgatgcg ggcggcgctg cagagccagc cgtccggcat 12120
taactcctcg gacgattgga cccaggccat gcaacgcatac atggcgctga cgacccgcaa 12180
tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt 12240
ggtgccctcg cgctcgaacc ccacgcacga gaagggtgctg gccatcgtga acgcgctggt 12300
ggagaacaag gccatccgcg gcgacgaggc cgggctggtg tacaacgcgc tgctggagcg 12360
cgtggcccg c tacaacagca ccaacgtgca gacgaacctg gaccgcatgg tgaccgacgt 12420
gcgcgaggcg gtgtcgcagc gcgagcggtt ccaccgcgag tcgaacctgg gctccatggt 12480
ggcgctgaac gccttcctga gcacgcagcc cgccaacgtg ccccggggccc aggaggacta 12540
caccaacttc atcagcgcgc tgcggctgat ggtggccgag gtgccccaga gcgaggtgta 12600
ccagtcgggg ccggactact tcttcagac cagtcgccag ggcttgca cgtgaacct 12660
gagccaggct ttcaagaact tgcagggact gtggggcgctg caggccccgg tcggggaccg 12720
cgcgacggtg tcgagcctgc tgacgccgaa ctgcgcctg ctgctgctgc tgggtggcgcc 12780
cttcacggac agcggcagcg tgagccgcga ctctacctg ggctacctgc ttaacctgta 12840
ccgcgaggcc atcgggcagg cgcacgtgga cgagcagacc taccaggaga tcacccacgt 12900
gagccgcgcg ctgggcccagg aggacccggg caacctggag gccaccctga acttcctgct 12960
gaccaaccgg tcgcagaaga tcccgcccca gtacgcgctg agcaccgagg aggagcgcat 13020
cctgcgctac gtgcagcaga gcgtggggct gttcctgatg caggaggggg ccacgcccag 13080
cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc 13140
gttcatcaat aagctgatgg actacttgca tcgggcggcc gccatgaact cggactactt 13200
taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta 13260
cgacatgccc gaccccaacg acgggttcct gtgggacgac gtggacagca gcgtgttctc 13320
gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggccgtcctc 13380
ggcgctgtcc ggtcgcgcg gtgctgccgc ggcggtgccc gaggccgcca gccccttccc 13440
gagcctgccc ttttcgctga acagcgtgcg cagcagcgag ctgggtcggc tgacgcggcc 13500
gcgcctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccc agcgcgagaa 13560
gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta 13620

cgcgacgag cacagggacg agccccgagc tagcagcagc accggcgcca cccgtagacg 13680
 ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacag 13740
 cagcgtgttg gacttgggtg ggagtgggtg tggttaacccg ttcgctcacc tgcgcccccg 13800
 tatcggggcg ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg 13860
 cgaccagcgt gcgttcttct ctgttggttg tagtagtatg atgaggcgcg tgtaccgga 13920
 gggtcctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcgggcg cgatgcagcc 13980
 cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag 14040
 cattcgttac tcggagctgg cacccttgta cgataccacc cggttgtagc tgggtggacaa 14100
 caagtcggcg gacatcgctt cgctgaacta ccagaacgac cacagcaact tcctgaccac 14160
 cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga 14220
 cgagcgctcg cggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280
 gaacgagttc atgtacagca acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc 14340
 caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt ggggtggagtt 14400
 tgagctgccc gagggcaact tctcgggtgac catgaccatc gatctgatga acaacgccat 14460
 catcgacaac tacttggcgg tggggcgcca gaacgggggtg ctggagagcg acatcggcgt 14520
 gaagttcgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tgggtgatgcc 14580
 gggcgtgtac accaacgagg ccttccaccc cgacatcgtc ctgctgcccg gctgcggcgt 14640
 ggacttcacc gagagccgcc tcagcaacct gctgggcacg cgcaagcggc agcccttcca 14700
 ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga 14760
 cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcggcgaccg cagccgtggc 14820
 caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcagcggc 14880
 cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga 14940
 gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc 15000
 ctacaactac ggcgaccccg agaaggcggt gcgctcctgg acgctgctca ccacctcgga 15060
 cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagacccggt 15120
 caccttccgc tccacgcgtc aagttagcaa ctaccgggtg gtgggcgccg agctcctgcc 15180
 cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt 15240
 cacctcgctc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgccccgc 15300
 cgcgcccacc attaccaccg tcagtgaaaa cgttcctgct ctacagatc acgggaccct 15360
 gccgctgcgc agcagtatcc ggggagtcga gcgcgtgacc gtcactgacg ccagacgccg 15420
 cacctgcccc tacgtctaca aggccctggg cgtagtcgcg ccgcgcgtcc tctcgagccg 15480
 caccttctaa aaaatgtcca ttctcatctc gccagtaat aacaccggtt ggggcctgcg 15540
 cgcgcccagc aagatgtacg gaggcgtcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600
 gcgcgggcac ttccgcgtc cctggggcg cctcaagggc cgcggtgcgt cgcgaccac 15660

cgctcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgccgccgc 15720
gcccgctctcc accgtggacg ccgtcatcga cagcgtggtg gccgacgcgc gccggtacgc 15780
ccgcgccaaag agccggcggc ggcgcacgc ccggcggcac cggagcacc ccgccatgcg 15840
cgcgggcgca gccttgctgc gcagggccag gcgcacggga cgcagggcca tgctcagggc 15900
ggccagacgc gcggcctccg gcagcagcag gccgggcagg accgcagac gcgcggccac 15960
ggcggcggcg gcggccatcg ccagcatgtc ccgccgcgg cgcggcaacg tgtactgggt 16020
gcgcgacgcc gccaccggtg tgcgcgtgcc cgtgcgcacc cccccctc gcacttgaag 16080
atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaattc 16140
aaggaagaga tgctccaggt catcgcgcct gagatctacg gcccggcggc ggtgaaggag 16200
gaaagaaagc ccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg 16260
gacggactgg tggagtttgt gcgcgagttc gcccccgcc ggcgctgca gtggcgcggg 16320
cggaagtga aaccggtgct gcgaccggc accacggtgg tcttcacgcc cggcgagcgt 16380
tccggctccg cctccaagcg ctctacgac gaggtgtacg gggacgagga catcctcgag 16440
caggcggccg aacgtctggg cgagtttgt tactggcaagc gcagccgcc cgcgcccttg 16500
aaagaggagg cgggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg 16560
gtgaccctgc agcaggtgct gcctggtgcg gcgcgcgcc ggggcttcaa gcgcgagggc 16620
ggcgaggatc tgtaccgac catgcagctg atggtgcca agcgccagaa gctggaggac 16680
gtgctggagc acatgaaggt ggaccccag gtgcagccc aggtcaaggt gcggcccatc 16740
aagcaggtgg ccccgggcct gggcggtcag accgtggaca tcaagatccc cacggagccc 16800
atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat 16860
ccctggatgc cggcaccggc ttccaccacc cgccgaagac gcaagtacgg cgcggccagc 16920
ctgctgatgc ccaactacgc gctgcatcct tccatcatcc ccacgccggg ctaccgcggc 16980
acgcgcttct accgcggcta caccagcagc cgccgccga agaccaccac ccgccgccgc 17040
cgctcgtcga cccgccgag cagcaccgcg acttccgccc cgccctggt gcggagagtg 17100
taccgcagcg ggcgcgagcc tctgaccctg ccgcgcgcgc gctaccacc gagcatcgcc 17160
atttaactac cgctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtcccat 17220
tacgggctac cgaggaagaa agccgcgccg tagaaggctg acggggaacg ggctgcgtcg 17280
ccatcaccac cggcggcggc gcgccatcag caagcgggtg gggggaggct tcctgcccgc 17340
gctgatgcc atcatcgcc cggcgatcgg ggcgatcccc ggcatagctt ccgtggcgg 17400
gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga 17460
cgctcctggt cctgtgatgt gtgttttttag atggaagaca tcaatttttc gtccctggca 17520
ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggcaa cagccaactg 17580
aacggggggc cttcaattg gagcagtctc tggagcgggc ttaagaattt cgggtccacg 17640
ctcaaacct atggcaacaa ggcgtggaac agcagcacag ggcaggcgct gagggaaaag 17700
ctgaaagagc agaacttcca gcagaaggtg gtcgatggcc tggcctcggg catcaacggg 17760

gtggtggacc tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc 17820
 ccgcccgcgg ggtccgtgga gatgccccag gtggaggagg agctgcctcc cctggacaag 17880
 cgcggcgaca agcgaccgcg tcccgcgcg gaggagacgc tgctgacgca cacggacgag 17940
 ccgccccgt acgaggaggc ggtgaaactg ggtctgcca ccacgcggcc cgtggcgcct 18000
 ctggccaccg ggggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg 18060
 cctccgcctg cttcccgcct ctccacagt gctaagcccc tgccgcgggt ggccgtcgcg 18120
 tcgcgcgccc cccgaggccg ccccaggcg aactggcaga gcactctgaa cagcatctg 18180
 ggtctgggag tgcagagtgt gaagcgccgc cgctgtatt aaaagacact gtagcgctta 18240
 acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgcgt 18300
 cgccgagttg caagatggcc accccatcga tgctgcccc gtgggcgtac atgcacatcg 18360
 ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag 18420
 acacctactt cagtctgggg aacaagttta ggaacccac ggtggcgccc acgcacgatg 18480
 tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccgtggac cgcgaggaca 18540
 acacctactc gtacaaagt cgctacacgc tggccgtggg cgacaaccgc gtgctggaca 18600
 tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaaccct 18660
 actccggcac cgcttacaac agcctggctc ccaagggagc gcccaacact tgccagtggg 18720
 catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgcctg 18780
 tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc 18840
 ccatttatgc agataaaact tatcaaccag agcctcaagt ggggtgatgct gaatggcatg 18900
 acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga 18960
 agccctgcta tggttctttt gccaaagccta ccaataaaga aggaggtcag gcaaagtga 19020
 aaaccgaaac aggcgggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa 19080
 gtgcagctgc ggctggcctg gcccagaaa ttgttttgta tactgagaat gtggatctgg 19140
 aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200
 atttgggtca gcagtccatg cccaacagac ccaactacat tggctttaga gacaacttta 19260
 tcgggctcat gtactacaac agcactggca acatgggcgt gctggctggg caggcctccc 19320
 agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc 19380
 ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct 19440
 atgaccccg tgtgcgcatt attgaaaatc acggtgtgga ggatgaactc cctaactatt 19500
 gcttccccct ggatgctgtg ggtagaactg atacttacca gggaattaag gccaatggtg 19560
 ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaataa ttgggcaagg 19620
 gcaatccttt cgccatggag atcaacatcc aggccaacct gtggcggaac ttcctctacg 19680
 cgaacgtggc gctgtacctg cccgactcct acaagtacac gccggccaac atcacgctgc 19740
 cgaccaacac caacacctac gattacatga acggccgcgt ggtggcgccc tcgctggtgg 19800

acgcctacat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaaccctt 19860
tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gtccttgggc aacgggcgct 19920
acgtgccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctcttgctcc 19980
tgcccgggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga 20040
gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc 20100
tctacgccac cttcttcccc atggcgcaca acaccgcctc cacgctcgag gccatgctgc 20160
gcaacgacac caacgaccag tccttcaacg actacctctc ggcggccaac atgctctacc 20220
ccatcccggc caacgccacc aacgtgcca tctccatccc ctgcgcgaac tgggccgcct 20280
tccgcggatg gtccttcacg cgcctcaaga cccgcgagac gccctcgctc ggctccgggt 20340
tcgaccctta cttcgtctac tcgggctcca tcccctacct cgacggcacc ttctacctca 20400
accacacctt caagaaggtc tccatcacct tcgactcctc cgtcagctgg cccggcaacg 20460
accgcctcct gacgccaac gagttcgaaa tcaagcgcac cgtcgacgga gagggggtaca 20520
acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gccactaca 20580
acatcggtta ccagggttc tacgtgcccg agggctacaa ggaccgcatg tactccttct 20640
tccgcaactt ccagcccatg agccgccagg tcgtggacga ggtcaactac aaggactacc 20700
aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctgcgcacca 20760
ccatgcgcca gggacagccc taccgcgcca actacccta cccgctcatc ggcaagagcg 20820
ccgtcgccag cgtcacccag aaaaagttcc tctgcgaccg ggtcatgtgg cgcattcccct 20880
tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg 20940
ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccaccc 21000
ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg 21060
tcatcgaggc cgtctacctg cgcacgccct tctcggccgg caacgccacc acctaagccc 21120
cgctcttgct tcttgcaaga tgacggcctg tgccggctcc ggcgagcagg agctcagggc 21180
catctccgc gacctgggct gcgggcccctg cttcctgggc accttcgaca agcgcttccc 21240
gggattcatg gccccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac 21300
cgggggcgag cactggctgg ctttcgcctg gaaccgcgc tcccacacct gctacctctt 21360
cgacccttc gggttctcgg acgagcgctt caagcagatc taccagttcg agtacgaggg 21420
cctgctgcgc cgcagcgcct tggccaccga ggaccgctgc gtcaccctgg aaaagtccac 21480
ccagaccgtg cagggtccgc gctcggccgc ctgcgggctc ttctgctgca tgttcctgca 21540
cgccttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgctgac 21600
gggggtgccc aacggcatgc tccagtcgcc ccaggtggaa cccaccctgc gccgcaacca 21660
ggaggcgctc taccgcttcc tcaacgcca ctccgcctac ttctgctccc accgcgcgcg 21720
catcgagaag gccaccgct tcgaccgat gaatcaagac atgtaaaccg tgtgtgtatg 21780
tgaatgcttt attcataata aacagcacat gtttatgcc acttttctga ggctctgact 21840
ttatttagaa atcgaagggg ttctgccggc tctcggcggtg cccgcgggc agggatacgt 21900

tgcggaactg gtacttgggc agccacttga actcggggat cagcagcttc ggcacgggga 21960
ggtcggggaa cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg 22020
gcgcggagat cttgaaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca 22080
cggggttgca gcaactggaac accatcaggg ccgggtgctt cacgctcgcc agcaccgctg 22140
cgtcggtgat gccctccacg tccagatcct cggcgttggc catcccgaag ggggtcatct 22200
tgcaggctctg ccgccccatg ctgggcacgc agccgggctt gtggttgcaa tcgcagtgca 22260
gggggatcag catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga 22320
aagcctccag ctggcggaag gcctgctcgg ccttgccgcc ctcggtgaag aagaccccgc 22380
aggacttgct agagaactgg ttggtggcgc agccggcgctc gtgcacgcag cagcgcgctg 22440
cgttgttggc cagctgcacc acgctgcgcc ccagcggtt ctgggtgatc ttggcccgg 22500
cggggttctc cttcagcgcg cgctgcccgt tctcgctcgc cacatccatc tcgatcgtgt 22560
gtccttctg gatcatcacg gtcccgtgca ggcatcgag cttgccctcg gcctcggtgc 22620
accctgtagc ccacagcgcg cagccggtgc actcccagtt cttgtgggcg atctgggagt 22680
gcgagtgac gaagccctgc aggaagcggc ccatcatcgt ggtcagggtc ttgttgctgg 22740
tgaaggtagc cgggatgccg cgggtgctct cgttcacata cagggtggcag atgcggcggt 22800
acacctcgcc ctgctcgggc atcagctgga aggcggactt caggctcgctc tccacgcggt 22860
accggtccat cagcagcgctc atgacttcca tgcccttctc ccaggccgag acgatcggca 22920
ggctcagggg gttcttcacc gccgttgtca tcttagtcgc cgccgctgag gtcaggggg 22980
cgttctcgctc cagggtctca aacactcgct tgccgtcctt ctcggtgatg cgcacggggg 23040
gaaagctgaa gcccacggcc gccagctcct cctcggcctg ctttctgtcc tcgctgtcct 23100
ggctgatgtc ttgcaaaggc acatgcttgg tcttgcgggg tttctttttg ggcggcagag 23160
gcggcgggcg agacgtgctg ggcgagcgcg agttctcgct caccacgact atttcttctt 23220
cttggccgtc gtccgagacc acgcggcggt aggcattgcct cttctggggc agaggcgag 23280
gcgacgggct ctgcgggttc ggcggggcggc tggcagagcc ctttcgcgt tcgggggtgc 23340
gtccttggcg gcgctgctct gactgacttc ctccgcggcc ggccattgtg ttctcctagg 23400
gagcaacaag catggagact cagccatcgt cgccaacatc gccatctgcc cccgccgccg 23460
ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gccgccagc cccacctcg 23520
acgccgccgc ggccccagac atgcaagaga tggaggaatc catcgagatt gacctgggct 23580
acgtgacgcc cgcgagcac gaggaggagc tggcagcgcg cttttcagcc ccggaagaga 23640
accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcagac 23700
atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcatctggcc cgccaatgca 23760
tcatcgtcaa ggacgcgtg ctgcaccgcg ccgaggtgcc cctcagcgtg gcggagctca 23820
gccgcgccta cgagcgcaac ctcttctcgc cgcgctgcc cccaagcgc cagcccaacg 23880
gcacctgcga gcccaaccg cgctcaact tctaccgggt cttcgcggtg cccgaggccc 23940

tggccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc 24000
gcacccgcgc cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatcgctt 24060
ccttggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggccgcga 24120
acgctctgca aggaagcgga gaggagcatg agcaccacag cgccctggtg gagttggaag 24180
gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgaccacac ttcgcctacc 24240
cggcgctcaa cctgcccccc aaggctcatga gcgccgtcat ggaccagggtg ctcatcaagc 24300
gcgcctcgcc cctctcggtat gaggacatgc aggaccccga gagctcggaac gagggcaagc 24360
ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gagtagcacc ccccagagct 24420
tggaagagcg gcgcaagctc atgatggccg tggctcctggt gaccgtggag ctggagtgtc 24480
tgcgccgctt cttcgccgac gcagagaccc tgcgcaaggt cgaggagaac ctgcactacc 24540
tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc 24600
tggctctcta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca 24660
ccctgcgcgg ggaggcccg cgcgactaca tccgcgactg cgtctacctg tacctctgcc 24720
acacctggca gacgggcatg ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag 24780
agctctgcaa gctcctgcag aagaacctga aggccctgtg gaccgggttc gacgagcgca 24840
ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca 24900
acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct ttcctcctcg 24960
aacgctccgg gatcctgccc gccacctgct ccgcgtgcc ctcggacttc gtgccgctga 25020
ccttccgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact 25080
acctggccta cactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140
actgccgctg caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga 25200
gcgagaccca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260
tgaaactcac cccggggctg tggacctcg cctacttgcg caagttcgtg cccgaggact 25320
accatccctt cgagatcagg ttctacgagg accaatccca gccgccaag gccgaactgt 25380
cggcctgcgt catcacccag ggggccatcc tggcccaatt gcaagccatc cagaaatccc 25440
gccagaatt tctgctgaaa aagggccacg ggggtctacct ggacccccag accggagagg 25500
agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag 25560
ctgccgccgc cggaggattt ggaggaagac tgggagagca gtcaggcaga ggaggaggag 25620
atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac 25680
gaggtggagg aggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctcggcggag 25740
aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcggcga ccgggcccac 25800
agtaggtggg acgagaccgg gcgcttcccg aaccccacca cccagaccgg taagaaggag 25860
cggcagggat acaagtcctg gcgggggcac aaaaacgcca tcgtctcctg cttgcaagcc 25920
tgcgggggca acatctcctt caccgcgcgc tacctgctct tccaccgcgg ggtgaacttc 25980
ccccgcaaca tcttgcatca ctaccgtcac ctccacagcc cctactactg tttccaagaa 26040

gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc 26100
ggcggcaggt ggactgagga tcgcagcgaa cgagccggcg cagacccggg agctgaggaa 26160
ccggatcttt cccaccctct atgccatctt ccagcagagt cgggggcagg agcaggaact 26220
gaaagtcaag aaccgtttct tgcgtctgct caccgcagct tgtctgtatc acaagagcga 26280
agaccaactt cagcgcactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgct 26340
cactcttaaa gagtagcccg cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc 26400
acctgcgccc ttgcgccgac catcatcatg agcaaagaga ttcccacgcc ttacatgtgg 26460
agctaccagc cccagatggg cctggccgcc ggcgccgcc aggactactc caccgcagct 26520
aactggctca gcgccgggccc cgcatgatc tcacgggtga atgacatccg cgcccgccga 26580
aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg 26640
cgtaattggc ccgccgccct ggtgtaccag gaaattcccc agcccacgac cgtactactt 26700
ccgcgagacg cccaggccga agtccagctg actaactcag gtgtccagct ggccggcggc 26760
gccgccctgt gtcgtcaccg ccccgctcag ggtataaagc ggctgggtgat ccgaggcaga 26820
ggcacacagc tcaacgacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc 26880
ttccaactcg ccggatcggg gagatcttcc ttcacgcctc gtcaggccgt cctgactttg 26940
gagagttcgt cctcgcagcc ccgctcgggt ggcacgcgca ctctccagtt cgtggaggag 27000
ttcactccct cggctctactt caacccttc tccggctccc ccggccacta cccggacgag 27060
ttcatcccga acttcgacgc catcagcgag tcggtggacg gctacgattg aatgtcccat 27120
ggtggcgag ctgacctagc tcggcttcga cacctggacc actgccgccg cttccgctgc 27180
ttcgctcggg atctcgccga gtttgccctac tttagctgc ccgaggagca ccctcagggc 27240
ccggcccacg gagtgcggat catcgtcgaa gggggcctcg actccacct gcttcggatc 27300
ttcagccagc gaccgatcct ggtcgagcgc gagcaaggac agacccttct gaccctgtac 27360
tgcacttgca accaccccgg cctgcatgaa agtctttgtt gtctgctgtg tactgagtat 27420
aataaaagct gagatcagcg actactccgg actcgattgt ggtgttcctg ctatcaaccg 27480
gtccctgttc ttaccggga acgagaccga gctccagctt cagtgtgaagc cccacaagaa 27540
gtacctcacc tggtgtttcc agggctcccc gatcgccgtt gtcaaccact gcgacaacga 27600
cggagtcctg ctgagcggcc ccgccaacct tactttttcc acccgagaa gcaagctcca 27660
gctcttccaa cccttcctcc ccgggacctc tcagtgcgtc tcgggaccct gccatcacac 27720
cttccacctg atcccgaata ccacagcgcc gctccccgct actaacaacc aaactacca 27780
ccatcgccac cgtcgcgacc tttctgaatc taacactacc acccacaccg gaggtgagct 27840
ccgaggtcga ccaacctctg ggatttacta cggcccctgg gaggtggtgg ggttaatagc 27900
gctaggccta gttgtgggtg ggcttttggc tctctgctac ctatacctcc cttgctgttc 27960
gtacttagtg gtgctgtgtt gctggtttaa gaaatgggga agatcacctt agtgagctgc 28020
ggtgctggtg tggcggtggt ggtgttttcg attgtgggac tgggcggcgc ggctgtagtg 28080

aaggagaagg cccgatccctg cttgcatttc aatccccgaca attgccagct gagttttcag 28140
cccgatggca atcgggtgcmc ggtgctgac aagtgcggat gggaatgcga gaacgtgaga 28200
atcgagtaca ataacaagac tcggaacaat actctcgcgt ccgtgtggca gcccggggac 28260
cccgagtggc acaccgtctc tgtccccggt gctgacggct ccccgcgac cgtgaacaat 28320
actttcattt ttgcmcacat gtgcmgacac gtcattgtga tgagcaagca gtacgatatg 28380
tgcccccca cgaaggagaa catcgtggc ttctccatcg cttacagcmc gtgcmggcg 28440
ctaaccaccg ctatcgtgtg cctgagcatt cacatgctca tcgctattcg cccagaaat 28500
aatgccgaaa aagagaaaca gccataacac gttttttcac acacctttt cagaccatgg 28560
cctctgttaa atttttgctt ttatttgcca gtctcattac tgttataagt aatgagaaac 28620
tcactattta cattggcact aaccacactt tagacggaat tccaaaatcc tcatggtatt 28680
gctattttga tcaagatcca gacttaacta tagaactgtg tggtacaag ggaaaaata 28740
caagcattca ttttaattaac ttttaattgcm gagacaattt gaaattaatt aatatcacta 28800
aagagtatgg aggtatgtat tactatgttg cagaaaataa caacatgcag ttttatgaag 28860
ttactgtaac taatcccacc acacctagaa caacaacaac caccaccaca aaaactacac 28920
ctgttaccac tatgcmgctc actaccaata acatttttgc catgcmgtcaa atggtcaaca 28980
atagcactca acccaccaca cccagtgagg aaattcccaa atccatgatt ggcattattg 29040
ttgctgtagt ggtgtgcmg ttgatcmg ccttggtgcm ggtgtactat gccttctgct 29100
acagaaagca cagactgaac gacaagctgg aacacttact aagtgttgaa ttttaatttt 29160
ttagaacat gaagatccta ggccttttaa ttttttctat cattacctt gctctatgca 29220
attctgacaa tgaggacgtt actgtcgttg tcggaaccaa ttataactg aaaggtccag 29280
cgaagggat gctttcgtgg tattgctgg ttggaactga cgagcaacag acagagctct 29340
gcaatgctca aaaaggcaaa acctcaaatt ctaaaatctc taattatcaa tgcaatggca 29400
ctgacttagt actgctcaat gtcacgaaag catatgctgg cagctacacc tgccctggag 29460
atgatactga gaacatgatt ttttacaag tggaagtgg ttgatccact actccacctc 29520
caccaccac aactactcac accacacaca cagaacaaac cacagcagag gaggcagcaa 29580
agttagcctt gcaggtccaa gacagttcat ttgttggcat taccctaca cctgatcagc 29640
ggtgtccggg gctgctcgtc agcggcattg tcggtgtgct ttcgggatta gcagtcataa 29700
tcatctgcat gttcattttt gcttgctgct atagaaggct ttaccgacaa aaatcagacc 29760
cactgctgaa cctctatgtt taattttttc cagagccatg aaggcagtta gcactctagt 29820
tttttgttct ttgattggca ctgttttttag tgttagcttt ttgaaacaaa tcaatgttac 29880
tgagggggaa aatgtgacac tggtaggcgt agagggtgct caaaatacca cctggacaaa 29940
attccatcta gatgggtgga aagaaatttg cacctggaat gtcagtactt atacatgtga 30000
aggagttaat cttaccattg tcaatgtcag ccaaattcaa aagggttgga ttaaagggca 30060
atctgttagt gttagcaata gtgggtacta taccagcat actcttatct atgacattat 30120
agttatacca ctgcctacac ctagcccacc tagcactacc acacagacaa cccacactac 30180

acaaacaacc acatacagta catcaaatca gcctaccacc actacaacag cagaggttgc 30240
 cagctcgtct ggggtccgag tggcattttt gatgttggcc ccatctagca gtccactgc 30300
 tagtaccaat gagcagacta ctgaattttt gtccactgtc gagagccaca ccacagctac 30360
 ctcgagtgcc ttctctagca ccgccaatct atcctcgctt tcctctacac caatcagtcc 30420
 cgctactact cctacccccg ctattctccc cactcccctg aagcaaacag acggcgacat 30480
 gcaatggcag atcacctgc tcattgtgat cgggttggtc atcctggccg tgttgctcta 30540
 ctacatcttc tgccgccgca ttcccaacgc gcaccgcaag ccggcctaca agcccatcgt 30600
 tgtcgggcag ccggagccgc ttcaggtgga agggggtcta aggaatcttc tcttctcttt 30660
 tacagtatgg tgattgaatt atgattccta gacaaatctt gatcactatt cttatctgcc 30720
 tcctccaagt ctgtgccacc ctgctctggg tggccaacgc cagtccagac tgtattgggc 30780
 ccttcgcctc ctacgtgctc tttgccttca tcacctgcat ctgctgctgt agcatagtct 30840
 gcctgcttat caccttcttc cagttcattg actggatctt tgtgcgcatc gcctacctgc 30900
 gccaccaccc ccagtaccgc gaccagcgag tggcgcggt gctcaggatc ctctgataag 30960
 catgcgggct ctgctacttc tcgcgcttct gctgttagtg ctccccgct ccgtcgaccc 31020
 ccggaccccc acccagtccc ccgaggaggt ccgcaaatgc aaattccaag aaccctggaa 31080
 attcctcaaa tgctaccgcc aaaaatcaga catgcatccc agctggatca tgatcattgg 31140
 gatcgtgaac attctggcct gcacctcat ctcttttggt atttaccctt gctttgactt 31200
 tggttggaac tcgccagagg cgctctatct cccgcctgaa cctgacacac caccacagca 31260
 acctcaggca cacgcactac caccaccacc acagcctagg ccacaataca tgcccatatt 31320
 agactatgag gccgagccac agcgacctat gctccccgct attagttact tcaatctaac 31380
 cggcgagat gactgacca ctggccaaca acaacgtcaa cgaccttctc ctggacatgg 31440
 acggccgcgc ctcgagcag cgactcgccc aacttcgcat tcgccagcag caggagagag 31500
 ccgtcaagga gctgcaggac ggcatagcca tccaccagtg caagaaaggc atcttctgcc 31560
 tggtgaaaca ggccaagatc tcctacgagg tcaccagac cgaccaatcg ctctcctacg 31620
 agctcctgca gcagcgccag aagttcacct gcctggtcgg agtcaacccc atcgtcatca 31680
 cccagcagtc gggcgatacc aaggggtgca tccactgctc ctgcgactcc cccgactgcg 31740
 tccacactct gatcaagacc ctctgcggcc tccgcgacct cctccccatg aactaatcac 31800
 ccccttatcc agtgaaataa agatcatatt gatgatttga gtttaataaa aataaagaat 31860
 cacttacttg aaatctgata ccaggctctt gtccatgttt tctgccaaca ccacttact 31920
 cccctcttcc cagctctggg actgcaggcc ccggcggtg gcaaacttcc tccacaccct 31980
 gaaggggatg tcaaattcct cctgtccctc aatcttcatt ttatcttcta tcagatgtcc 32040
 aaaaagcgcg tccgggtgga tgatgacttc gaccccgctt acccctacga tgcagacaac 32100
 gcaccgaccg tgcccttcat caaccccccc ttctgtctct cagatggatt ccaagagaag 32160
 cccctggggg tgctgtccct gcgtctggcc gatcccgctc ccaccaagaa cggggaaatc 32220

accctcaagc tgggagatgg ggtggacctc gactcctcgg gaaaactcat ctccaacacg 32280
 gccaccaagg ccgcccggcc tctcagtttt tccaacaaca ccatttcctt taacatggat 32340
 accccttttt acaacaacaa tggaaagtta ggcattgaaag tcaactgctcc actgaagata 32400
 ctagacacag acttgctaaa aacacttggt gtagcttatg gacaagggtt aggaacaaac 32460
 accactgggtg ccttggttgc ccaactagca tccccacttg cttttgatag caatagcaaa 32520
 attgccctta atttaggcaa tggaccattg aaagtggatg caaatagact gaacatcaat 32580
 tgcaatagag gactctatgt tactaccaca aaagatgcac tggaagccaa tataagttgg 32640
 gctaattgcta tgacatttat aggaaatgcc atgggtgtca atattgatac aaaaaaggc 32700
 ttgcaatttg gcaccactag taccgtcgca gatgttaaaa acgcttacc catacaaatc 32760
 aaacttggag ctggtctcac atttgacagc acaggtgcaa ttgttgcatg gaacaaagat 32820
 gatgacaagc ttacactatg gaccacagcc gaccctctc caaattgtca catatattct 32880
 gaaaaggatg ctaagcttac actttgcttg acaaagtgtg gcagtcagat tctgggcact 32940
 gtttcctca tagctgttga tactggcagt ttaaattcca taacaggaac agtaaccact 33000
 gctcttgtct cacttaaat cgatgcaaat ggagttttgc aaagcagctc aacactagac 33060
 tcagactatt ggaatttcag acagggagat gttacacctg ctgaagccta tactaatgct 33120
 ataggtttca tgcccaatct aaaagcatac cctaaaaaca caagtggagc tgcaaaaagt 33180
 cacattgttg ggaaagtgt cctacatggg gatacaggca aaccactgga cctcattatt 33240
 actttcaatg aaacaagtga tgaatcttgc acttactgta ttaactttca atggcagtgg 33300
 ggggctgatc aatataaaaa tgaaacactt gccgtcagtt cattcacctt ttcctatatt 33360
 gctaaagaat aaacccact ctgtacccca tctctgtcta tggaaaaaac tctgaaacac 33420
 aaaataaaat aaagttcaag tgttttattg attcaacagt ttacaggat tcgagcagtt 33480
 atttttcctc caccctcca ggacatggaa tacaccacc tctcccccg cacagccttg 33540
 aacatctgaa tgccattggg gatggacatg cttttggtct ccacgttcca cacagtttca 33600
 gagcgagcca gtctcgggtc ggtcaggag atgaaacct ccgggcactc ccgcatctgc 33660
 acctcacagc tcaacagctg aggattgtcc tcggtggtcg ggatcacggt tatctggaag 33720
 aagcagaaga gcggcggttg gaatcatagt ccgcgaacgg gatcggccgg tgggtgtcgca 33780
 tcaggccccg cagcagtcgc tgtcgccgcc gctccgtcaa gctgctgctc agggggtccg 33840
 ggtccaggga ctccctcagc atgatgcca cgccctcag catcagtcgt ctggtgcggc 33900
 gggcgagca gcgcatgcg atctcgctca ggtcgctgca gtacgtgcaa cacaggacca 33960
 ccaggttgtt caacagtcca tagttcaaca cgctccagcc gaaactcatc gcgggaagga 34020
 tgctaccac gtggccgtcg taccagatcc tcaggtaaata caagtggcg cccctccaga 34080
 acacgctgcc catgtacatg atctccttgg gcatgtggcg gttcaccacc tcccgttacc 34140
 acatcacct ctggttgaac atgcagcccc ggatgatcct gcggaaccac agggccagca 34200
 ccgccccgcc cgccatgcag cgaagagacc ccgggtcccg acaatggcaa tggaggaccc 34260
 accgctcgta ccggtggatc atctgggagc tgaacaagtc tatgttggca cagcacaggc 34320

atatgctcat	gcattctcttc	agcactctca	gctcctcggg	ggtcaaaacc	atatcccagg	34380
gcacggggaa	ctcttgacag	acagcgaacc	ccgcagaaca	gggcaatcct	cgcacataac	34440
ttacattgtg	catggacagg	gtatcgcaat	caggcagcac	cggtgatcc	tccaccagag	34500
aagcgcgggt	ctcgggtctcc	tcacagcgtg	gtaagggggc	cggccgatac	gggtgatggc	34560
gggacgcggc	tgatcgtgtt	cgcgaccgtg	ttatgatgca	gttgctttcg	gacattttcg	34620
tacttgctgt	agcagaacct	ggtccggggc	ctgcacaccg	atcgccggcg	gcggtcccgg	34680
cgtttggaac	gctcgggtgtt	gaagttgtaa	aacagccact	ctctcagacc	gtgcagcaga	34740
tctagggcct	caggagtgat	gaagatccca	tcatgcctga	tggctctaata	cacatcgacc	34800
accgtggaat	gggccagacc	cagccagatg	atgcaatttt	gttggggtttc	ggtgacggcg	34860
ggggagggaa	gaacaggaag	aaccatgatt	aacttttaata	ccaaacgggtc	tcggagcact	34920
tcaaatgaa	gatcgcgag	atggcacctc	tcgccccgc	tgtgttggtg	gaaaataaca	34980
gccaggtcaa	aggtgatacg	gttctcgaga	tgttccacgg	tggcttcag	caaagcctcc	35040
acgcgcacat	ccagaaacaa	gacaatagcg	aaagcgggag	ggttctctaa	ttcctcaatc	35100
atcatgttac	actcctgcac	catccccaga	taattttcat	ttttccagcc	ttgaatgatt	35160
cgaactagtt	cctgaggtaa	atccaagcca	gccatgataa	agagctcgcg	cagagcggcc	35220
tccaccggca	ttcttaagca	caccctcata	attccaagat	attctgctcc	tggttcacct	35280
gcagcagatt	gacaagcggg	atatcaaaat	ctctgccgcg	atccctaagc	tcctccctca	35340
gcaataactg	taagtactct	ttcatatcct	ctccgaaatt	tttagccata	ggaccaccag	35400
gaataagatt	agggcaagcc	acagtacaga	taaaccgaag	tcctccccag	tgagcattgc	35460
caaatgcaag	actgctataa	gcattgctggc	tagaccgggt	gatattcttc	agataactgg	35520
acagaaaatc	gcccaggcaa	tttttaagaa	aatcaacaaa	agaaaaatcc	tccaggtgca	35580
cgttttagagc	ctcgggaaca	acgatggagt	aatgcaagc	ggtgcgttcc	agcatggtta	35640
gttagctgat	ctgtagaaaa	aaacaaaaat	gaacattaaa	ccattgctagc	ctggcgaaca	35700
ggtgggtaaa	tcgttctctc	cagcaccagg	caggccacgg	ggtctccggc	acgaccctcg	35760
taaaaattgt	cgctatgatt	gaaaaccatc	acagagagac	gttcccgggtg	gccggcgtga	35820
atgattcgac	aagatgaata	cacccccgga	acattggcgt	ccgcgagtga	aaaaaagcgc	35880
ccaaggaagc	aataaggcac	tacaatgctc	agtctcaagt	ccagcaaagc	gatgccatgc	35940
ggatgaagca	caaaattctc	aggtgcgtac	aaaatgtaat	tactcccctc	ctgcacaggc	36000
agcaaagccc	ccgatccctc	caggtaacaa	tacaaagcct	cagcgtccat	agcttaccga	36060
gcagcagcac	acaacaggcg	caagagtcag	agaaaggctg	agctctaacc	tgtccacccg	36120
ctctctgctc	aatatatagc	ccagatctac	actgacgtaa	aggccaaagt	ctaaaaatac	36180
ccgccaaata	atcacacacg	cccagcacac	gcccagaaac	cggtgacaca	ctcaaaaaaa	36240
tacgcgcact	tcctcaaacg	cccaaactgc	cgtcatttcc	gggttcccac	gctacgtcat	36300
caaaattcga	ctttcaaatt	ccgtcgaccg	ttaaaaacgt	cgcccgcccc	gcccctaacg	36360

gtcgcgctc ccgcagccaa tcaccgcccc gcatccccaa attcaaatac ctcatTTgca 36420
 tattaacgcg caccaaaagt ttgaggtata ttattgatga tg 36462

<210> 2
 <211> 36604
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan6

<400> 2
 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaata agctgtttga 60
 atttggggag ggaggaaggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
 tgacgttttg atgacgtggc tatgagggcg agccggtttg caagtTctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaataactcaa ttttcccgcg ctctctgaca 240
 ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggta 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagtTttct 540
 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
 gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtgggtggac gccatgatgg 660
 gtgacgaccc tccagagccc cctaccccat ttgagggccc ttcgctgtac gatttgtatg 720
 atctggaggt ggatgtgccc gagagcgacc ctaacgagga ggcggtgaat gatttgttta 780
 gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
 cctctctcca taccgccaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgcg gtgaaccagg gagtgaaaac tgcgggagag agcttttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
 acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactggttt 1200
 atttatgtat atgttttttt atgtgtaggt cccgtctctg acgtagatga gacccccact 1260
 tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
 agaccagttg cagtgaagat caccgggagg agagcagctg tggagagttt ggatgacttg 1380
 ctacaggggt gggatgaacc tttggacttg tgtaccgga aacgccccag gcactaagtg 1440
 ccacacatgt gtgtttactt aaggtgatgt cagtatttat aggggtgtgga gtgcaataaa 1500
 atccgtgttg actttaagtg cgtgttttat gactcagggg tggggactgt gggatatataa 1560
 gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gactgtcttg 1620
 gaagactttc accagactag acagttgcta gagaactcat cggagggagt ctcttacctg 1680
 tggagattct gcttcgggtg gcccttagct aagctagtct atagggccaa acaggattat 1740

aaggaacaat	ttgaggatat	tttgagagag	tgtcctggta	tttttgactc	tctcaacttg	1800
ggccatcagt	ctcactttta	ccagagtatt	ctgagagccc	ttgacttttc	tactcctggc	1860
agaactaccg	ccgcggtagc	cttttttgcc	tttattcttg	acaaatggag	tcaagaaacc	1920
catttcagca	gggattaccg	tctggactgc	ttagcagtag	ctttgtggag	aacatggagg	1980
tgccagcgcc	tgaatgcaat	ctccggctac	ttgccagtag	agccggtaga	cacgctgagg	2040
atcctgagtc	tccagtcacc	ccaggaacac	caacgccgcc	agcagccgca	gcaggagcag	2100
cagcaagagg	aggaccgaga	agagaacccg	agagccggtc	tggaccctcc	ggtggcggag	2160
gaggaggagt	agctgacttg	tttcccagac	tgcgccgggt	gctgactagg	tcttccagtg	2220
gacgggagag	ggggattaag	cgggagaggc	atgaggagac	tagccacaga	actgaactga	2280
ctgtcagtct	gatgagccgc	aggcgccag	aatcggtgtg	gtggcatgag	gtgcagtcgc	2340
aggggataga	tgaggtctcg	gtgatgcatg	agaaatattc	cctagaacaa	gtcaagactt	2400
gttggttga	gcccaggat	gattgggagg	tagccatcag	gaattatgcc	aagctggctc	2460
tgaagccaga	caagaagtac	aagattacca	aactgattaa	tatcagaaat	tcctgctaca	2520
tttcaggga	tggggccgag	gtggagatca	gtaccagga	gagggtgcc	ttcagatgtt	2580
gtatgatgaa	tatgtacccg	ggggtggtgg	gcatggaggg	agtcaccttt	atgaacacga	2640
ggttcagggg	tgatgggtat	aatgggggtg	tctttatggc	caacaccaag	ctgacagtgc	2700
acggatgctc	cttctttggc	ttcaataaca	tgtgcatcga	ggcctggggc	agtgtttcag	2760
tgaggggatg	cagcttttca	gccaactgga	tgggggtcgt	gggcagaacc	aagagcaagg	2820
tgtcagtga	gaaatgcctg	ttcgagaggt	gccacctggg	ggtgatgagc	gagggcgaag	2880
ccaaagtcaa	acactgcgcc	tctaccgaga	cgggctgctt	tgtgctgata	aagggcaatg	2940
cccaagtcaa	gcataacatg	atctgtgggg	cctcgatga	gcgcggctac	cagatgctga	3000
cctgcgccgg	tgggaacagc	catatgctgg	ccacctgca	tgtggcctcg	cacccccgca	3060
agacatggcc	cgagttcgag	cacaacgtca	tgacctgctg	caatgtgcac	ctgggctccc	3120
gccgaggcat	gttcatgccc	taccagtga	acatgcaatt	tgtgaagggtg	ctgctggagc	3180
ccgatgccat	gtccagagtg	agcctgacgg	gggtgtttga	catgaatgtg	gagctgtgga	3240
aaattctgag	atatgatgaa	tccaagacca	ggtgccgggc	ctgcgaatgc	ggaggcaagc	3300
acgccaggct	tcagcccgtg	tgtgtggagg	tgacggagga	cctgcgaccc	gatcatttgg	3360
tgttgtcctg	caacgggacg	gagttcggct	ccagcgggga	agaatctgac	tagagttagt	3420
agtgtttggg	gctgggtgtg	agcctgcatg	aggggcagaa	tgactaaaat	ctgtggtttt	3480
ctgtgtgttg	cagcagcatg	agcggaagcg	cctcctttga	gggaggggta	ttcagccctt	3540
atctgacggg	gcgtctcccc	tcctgggcgg	gagtgcgtca	gaatgtgatg	ggatccacgg	3600
tggacggccg	gcccgtgcag	cccgcgaact	cttcaaccct	gacctacgcg	accctgagct	3660
cctcgtccgt	ggacgcagct	gccgccgcag	ctgctgcttc	cgccgccagc	gccgtgcgcg	3720
gaatggccct	gggcgccggc	tactacagct	ctctggtggc	caactcgagt	tccaccaata	3780
atcccgcag	cctgaacgag	gagaagctgc	tgctgctgat	ggcccagctc	gaggccctga	3840

```

cccagcgcct gggcgagctg acccagcagg tggctcagct gcaggcggag acgcgggccc 3900
cggttgccac ggtgaaaacc aaataaaaaa tgaatcaata aataaacgga gacggttggt 3960
gattttaaca cagagtcttg aatctttatt tgatttttcg cgcgcggtag gccctggacc 4020
accggtctcg atcattgagc acccggtgga tcttttccag gacccggtag aggtgggctt 4080
ggatgttgag gtacatgggc atgagcccgt cccgggggtg gaggtagctc cattgcaggg 4140
cctcgtgctc ggggatgggtg ttgtaaataca cccagtcata gcagggggcgc agggcgtgggt 4200
gctgcacgat gtccttgagg aggagactga tggccacggg cagccccttg gtgtagggtg 4260
tgacgaacct gttgagctgg gagggatgca tgcgggggga gatgagatgc atcttggcct 4320
ggatcttgag attggcgatg ttcccgccca gatcccgcg ggggttcatg ttgtgcagga 4380
ccaccagcac ggtgtatccg gtgcacttg ggaatttgtc atgcaacttg gaagggaagg 4440
cgtgaaagaa tttggagacg cccttgtgac cggccagggtt ttccatgcac tcatccatga 4500
tgatggcgat gggcccgtgg gcggcggcct gggcaaagac gtttcggggg tcggacacat 4560
cgtagttgtg gtcctgggtg agctcgtcat aggccatttt aatgaatttg gggcggaggg 4620
tgcccactg ggggacgaag gtgccctcga tcccgggggc gtagttgcc tcgcagatct 4680
gcatctccca ggccttgagc tcggaggggg ggatcatgtc cacctgcggg gcgatgaaaa 4740
aaacggtttc cggggcgggg gagatgagct gggccgaaag caggttccgg agcagctggg 4800
acttgccgca accggtgggg ccgtagatga ccccgatgac cggctgcagg tggtagttga 4860
gggagagaca gctgccgtcc tcgaggagga gggggggccac ctcgttcatc atctcgcgca 4920
catgcatgtt ctgcgcacg agttccgcca ggaggcgctc gccccccagc gagaggagct 4980
cttgacgca ggcgaagttt ttcagcggct tgagtccgtc ggccatgggc attttggaga 5040
gggtctgttg caagagttcc agacgggtcc agagctcggg gatgtgctct agggcatctc 5100
gatccagcag acctcctcgt ttcgcgggtt ggggcgactg cgggagtagg gcaccaggcg 5160
atgggcgtcc agcaggcca gggtcgggtc cttccagggc cgcagggtcc gcgtcagcgt 5220
gggtctcgtc acggtgaagg ggtgcgcgcc gggctgggag cttgcgaggg tgcgcttcag 5280
gctcatccgg ctggtcgaga accgctccc gtcggcgccc tgcgcgtcgg ccaggtagca 5340
attgagcatg agttcgtagt tgagcgctc ggccgcgtgg cccttggcgc ggagcttacc 5400
tttgaagtg tgtccgcaga cgggacagag gagggacttg agggcgtaga gcttgggggc 5460
gaggaagacg gactcggggg cgtaggcgctc cgcgccgag ctggcgcaga cgggtctcgca 5520
ctccacgagc caggtgaggt cggggcggtt ggggtcaaaa acgaggtttc ctccgtgctt 5580
tttgatgcgt ttcttacctc tgggtctccat gagctcgtgt ccccgctggg tgacaaagag 5640
gctgtccgtg tccccgtaga ccgactttat gggccgggtc tcgagcgggg tgccgcggtc 5700
ctcgtcgtag aggaacccc cccactccga gacgaaggcc cgggtccagg ccagcacgaa 5760
ggaggccacg tgggaggggt agcggtcgtt gtccaccagc ggggtccacct tctccagggt 5820
atgcaagcac atgtccccct cgtccacatc caggaagggt attggcttgt aagtgtaggc 5880

```

cacgtgaccg ggggtccccg ccgggggggt ataaaagggg gcggggcccct gctcgtcctc 5940
actgtcttcc ggatcgctgt ccaggagcgc cagctgttgg ggtaggtatt ccctctcgaa 6000
ggcgggcatg acctcggcac tcaggttgtc agtttctaga aacgaggagg atttgatatt 6060
gacggtgccg ttggagacgc ctttcatgag cccctcgtcc atttggtcag aaaagacgat 6120
ctttttgttg tcgagcttgg tggcgaagga gccgtagagg gcgttggaga gcagcttggc 6180
gatggagcgc atggtctggt tcttttcctt gtcggcgcg ctccttggcg cgatgttgag 6240
ctgcacgtac tcgcgcgcca cgcacttcca ttcggggaag acggtggtga gctcgtcggg 6300
cacgattctg acccgccagc cgcggttgtg cagggtgatg aggtccacgc tgggtggccac 6360
ctcgccgcgc aggggctcgt tgggtccagca gaggcgcccg cccttgcgcg agcagaaggg 6420
gggcagcggg tccagcatga gctcgtcggg ggggtcggcg tccacggtga agatgccggg 6480
caggagctcg gggtcgaagt agctgatgca ggtgcccaga ttgtccagcg ccgcttgcca 6540
gtcgcgcacg gccagcgcg gctcgtaggg gctgaggggc gtgccccagg gcatgggggtg 6600
cgtgagcgcg gaggcgatca tgccgcagat gtcgtagacg tagaggggct cctcaggagc 6660
gccgatgtag gtggggtagc agcgcccccc gcggtgctg gcgcgcacgt agtcgtacag 6720
ctcgtgagag ggcgcgagga gcccctgcc gaggttggag cgttgcggt tttcggcgcg 6780
gtagacgatc tggcggaaga tggcggtgga gttggaggag atggtgggccc tttggaagat 6840
gttgaagtgg gcgtggggca ggccgaccga gtccctgatg aagtgggctg aggagtcctg 6900
cagcttggcg acgagctcgg cggtgacgag gacgtccagg gcgcagtagt cgagggtctc 6960
ttggatgatg tcatacttga gctggccctt ctgcttccac agctcgcggt tgagaaggaa 7020
ctcttcgcgg tccttccagt actcttcgag ggggaacccg tcctgatcgg cacggtaaga 7080
gcccaccatg tagaactggt tgacggcctt gtaggcgcag cagcccttct ccacggggag 7140
ggcgtaagct tgcgcgccct tgcgcgaggga ggtgtgggtg agggcggaagg tgtcgcgcac 7200
catgaccttg aggaactggt gcttgaagtc gaggtcgtcg cagccgccct gctcccagag 7260
ttggaagtcc gtgcgcttct tgtaggcggg gttaggcaaa gcgaaagtaa catcgttgaa 7320
gaggatcttg ccgcgcggg gcatgaagtt gcgagtgatg cggaagggt ggggcacctc 7380
ggcccgggtt ttgatgacct gggcggcgag gacgatctcg tcgaagccgt tgatgttgtg 7440
cccgcgatg tagagttcca cgaatcgcgg gcggcccttg acgtggggca gcttcttgag 7500
ctcgtcgtag gtgagctcgg cggggtcgct gagcccgtgc tgctcagagg ccagtcggc 7560
gacgtggggg ttggcgctga ggaagggaagt ccagagatcc acggccaggg cggctctgaa 7620
gcggtccccg tactgacgga actgttggcc cacggccatt ttttcggggg tgacgcagta 7680
gaaggtgcgg gggctcgccgt gccagcggtc ccacttgagc tggagggcga ggtcgtgggc 7740
gagctcgacg agcggcgggg ccccgagag tttcatgacc agcatgaagg ggacgagctg 7800
cttgccgaag gaccccatcc aggtgtaggt ttccacatcg taggtgagga agagcctttc 7860
ggtgcgagga tgcgagccga tggggaagaa ctggatctcc tgccaccagt tggaggaatg 7920
gctgttgatg tgatggaagt agaaatgccg acggcgcgcc gagcactcgt gcttgtgttt 7980

atacaagcgt ccgcagtgct cgcaacgctg cacgggatgc acgtgctgca cgagctgtac 8040
ctgggttcct ttggcgagga atttcagtg gacagtgagc gctggcggct gcatctcgtg 8100
ctgtactacg tcttggccat cggcgtggcc atcgtctgcc tcgatggtgg tcatgtgac 8160
gagcccgcgc gggaggcagg tccagacctc ggctcggacg ggctggagag cgaggacgag 8220
ggcgcgacag ccggagctgt ccagggtcct gagacgctgc ggagtcaggt cagtgggcag 8280
cggcggcgcg cggttgactt gcaggagctt ttccagggcg cgcgggagggt ccagatggta 8340
cttgatctcc acggcgccgt tgggtggctac gtccacggct tgcagggtgc cgtgcccctg 8400
gggcgccacc accgtgcccc gtttcttctt gggcgctgct tccatgtcgg tcagaagcgg 8460
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg 8520
ggcacgtcgg cgccgcgcgc gggcagggtt tgggtactgcg cccggagaag actggcgtga 8580
gcgacgacgc gacggttgac gtcctggatc tgacgcctct ggggtgaaggc cacgggaccc 8640
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc 8700
tgccgcagga tctcttgac gtcgcccag ttgtcctggt aggcgatctc ggtcatgaac 8760
tgctcgatct cctcctcctg aaggctctccg cgcccggcgc gctcgacggt ggccgcgagg 8820
tcgttgagga tgcgcccat gagctgcgag aaggcggtca tgccggcctc gttccagacg 8880
cggctgtaga ccacggctcc gtcggggtcg cgcgcgcgca tgaccacctg ggcgagggtg 8940
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc 9000
gtggtggcga tgtgctcggg gacgaagaag tacatgatcc agcggcgag cgcatctcg 9060
ctgacgtcgc ccagggttc caagcggttc atggcctcgt agaagtccac ggcgaagttg 9120
aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg 9180
gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc catctcctcc 9240
tcttcctcct cactaacat ctcttctact tcctcctcag gaggcggtgg cgggggaggg 9300
gccctgcgtc gccggcggcg cacgggcaga cggctgatga agcgtcgat ggtctccccg 9360
cgccggcgac gcatggtctc ggtgacggcg cgcccgctc cgcggggccc cagcatgaag 9420
acgccgccgc gcatctccag gtggccgccg ggggggtctc cgttgggcag ggagagggcg 9480
ctgacgatgc atcttatcaa ttgacctgta gggactccgc gcaaggacct gagcgtctcg 9540
agatccacgg gatccgaaaa ccgctgaacg aaggcttcga gccagtcgca gtcgcaaggt 9600
aggctgagcc cggtttcttg ttcttcgggt atttggtcgg gaggcgggcg ggcgatgctg 9660
ctggtgatga agttgaagta ggcggtcctg agacggcgga tgggtggcgag gagcaccagg 9720
tccttggggc cggttgctg gatgacgaga cggctggcca tgcccaggc gtggtcctga 9780
cacctggcga ggtccttgta gtagtctgc atgagccgct ccacgggcac ctctcctcg 9840
cccgcgcggc cgtgcatgcg cgtgagcccc aaccgcgct gcggctggac gagcgccagg 9900
tcggcgacga cgcgctcggg gaggatggcc tgctggatct ggggtgagggt ggtctggaag 9960
tcgtcgaagt cgacgaagcg gtggtaggct ccggtgttga tgggttagga gcagttggcc 10020

atgacggacc agttgacggt ctggtggccg ggtcgcacga gctcgtggta cttgaggcgc 10080
gagtaggcgc gcgtgtcgaa gatgtagtcg ttgcaggcgc gcacgaggta ctggtatccg 10140
acgaggaagt gcggcgccg ctggcggtag agcggccatc gctcgggtggc gggggcgccg 10200
ggcgcgaggt cctcgagcat gaggcggtgg tagccgtaga tgtacctgga catccagggtg 10260
atgccggcgg cgggtggtgga ggcgcgcggg aactcgcgga cgcggttcca gatgttgccg 10320
agcggcagga agtagttcat ggtggccgcg gtctggcccc tgaggcgcgc gcagtcgtgg 10380
atgctctaga catacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag 10440
gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag 10500
ccgcagctaa cgtggtactg gcaactccgt ctcgacccaa gcctgctaac gaaacctcca 10560
ggatacggag gcgggtcggt ttttggcctt ggtcgtgggt catgaaaaac tagtaagcgc 10620
ggaaagcggc cgcccgcat ggctcgtgc cgtagtctgg agaaagaatc gccagggttg 10680
cgttgcggtg tgccccggtt cgagcctcag cgctcggcgc cggccggatt ccgcggctaa 10740
cgtgggcgtg gctgccccgt cgtttccaag accccttagc cagccgactt ctccagttac 10800
ggagcgagcc cctctttttt tttcttgtgt ttttgccaga tgcattccgt actgcggcag 10860
atgcgcccc accctccacc acaaccgccc ctaccgcagc agcagcaaca gccggcgctt 10920
ctgccccgc cccagcagca gccagccact accgcggcgg ccgccgtgag cggagccggc 10980
gttcagtatg acctggcctt ggaagagggc gaggggctgg cgcggctggg ggcgctcgtc 11040
ccggagcggc acccgcgctg gcagatgaaa agggacgctc gcgaggccta cgtgccaag 11100
cagaacctgt tcagagacag gagcggcgag gagcccagg agatgcgcgc ctcccgttc 11160
cacgcggggc gggagctgc gcgcggcctg gaccgaaagc ggggtgctgag ggacgaggat 11220
ttcaggcgcg acgagctgac ggggatcagc cccgcgcgcg cgcacgtggc cgcggccaac 11280
ctggtcacgg cgtacgagca gaccgtgaag gaggagagca acttcaaaa atccttcaac 11340
aaccacgtgc gcacgctgat cgcgcgcgag gaggtgacct tgggcctgat gcacctgtgg 11400
gacctgctgg aggccatcgt gcagaacccc acgagcaagc cgctgacggc gcagctgttt 11460
ctggtggtgc agcacagtcg ggacaacgag acgttcaggg aggcgctgct gaatatcacc 11520
gagcccagg gccgctggct cctggacctg gtgaacattt tgacagcat cgtggtgcag 11580
gagcggggc tgccgctgtc cgagaagctg gcggccatca acttctcggg gctgagtctg 11640
ggcaagtact acgctaggaa gatctacaag acccgtacg tgcccataga caaggagggtg 11700
aagatcgacg ggttttacat gcgcatgacc ctgaaagtgc tgaccctgag cgacgatctg 11760
ggggtgtacc gcaacgacag gatgcaccgc gcggtgagcg ccagccgccg gcgcgagctg 11820
agcgaccagg agctgatgca cagcctgcag cggggccctga ccggggccgg gaccgagggg 11880
gagagctact ttgacatggg cgcggacctg cgctggcagc ccagccgccg ggccttgga 11940
gctgccggcg gttcccccta cgtggaggag gtggacgatg aggaggagga gggcgagtac 12000
ctggaagact gatggcgcga ccgtatTTTT gctagatgca gcaacagcca ccgccgccgc 12060
ctcctgatcc cgcgatgcgg gcggcgctgc agagccagcc gtccggcatt aactcctcgg 12120

acgattggac ccaggccatg caacgcatca tggcgctgac gacccgcaat cccgaagcct 12180
ttagacagca gcctcaggcc aaccggctct cggccatcct ggaggccgtg gtgccctcgc 12240
gctcgaaccc cacgcacgag aagggtgctg ccatacgtgaa cgcgctggtg gagaacaagg 12300
ccatccgcgg tgacgaggcc gggctggtgt acaacgcgct gctggagcgc gtggcccgcct 12360
acaacagcac caacgtgcag acgaacctgg accgcatggt gaccgacgtg cgcgaggcgg 12420
tgtcgcagcg cgagcgggtt caccgcgagt cgaacctggg ctccatggtg gcgctgaacg 12480
ccttcctgag cacgcagccc gccaacgtgc cccggggcca ggaggactac accaacttca 12540
tcagcgcgct gcggctgatg gtggccgagg tgccccagag cgaggtgtac cagtcggggc 12600
cggactactt cttccagacc agtcgccagg gcttgacagc cgtgaacctg agccaggctt 12660
tcaagaactt gcagggactg tggggcgctg agggccccgg cggggaccgc gcgacggtgt 12720
cgagcctgct gacgccgaac tcgcgcctgc tgctgctgct ggtggcgccc ttcacggaca 12780
gcggcagcgt gagccgcgac tcgtacctgg gctacctgct taacctgtac cgcgaggcca 12840
tcggacaggc gcacgtggac gagcagacct accaggagat caccacgtg agccgcgcgc 12900
tgggccagga ggacccgggc aacctggagg ccacctgaa cttcctgctg accaaccggt 12960
cgcagaagat cccgccccag tacgcgctga gcaccgagga ggagcgcata ctgcgctacg 13020
tgcagcagag cgtggggctg ttcctgatgc aggagggggc cacgcccagc gcggcgctcg 13080
acatgaccgc gcgcaacatg gagcccagca tgtacgcccg caaccgcccg ttcataata 13140
agctgatgga ctacttgcat cgggcggccg ccatgaactc ggactacttt accaacgcca 13200
tcttgaaccc gcactggctc ccgccgcccg ggttctacac gggcgagtac gacatgcccg 13260
accccaacga cgggttcctg tgggacgacg tggacagcag cgtgttctcg ccgcgtccag 13320
gaaccaatgc cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg 13380
gtcgcgcggg tgctgccgcg gcggtgcccg agggcgccag ccccttcccg agcctgccct 13440
tttcgctgaa cagcgtgcgc agcagcgagc tgggtcggct gacgcgaccg cgcctgctgg 13500
gcgaggagga gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttcccca 13560
ataacgggat agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc 13620
acagggacga gccccgagct agcagcgagc gcacccgtag acgacagcgg cacgacaggc 13680
agcggggact ggtgtgggac gatgaggatt ccgccgacga cagcagcgtg ttggacttgg 13740
gtgggagtg tggtaacccg ttcgctcacc tgcgccccg tatcgggcgc ctgatgtaag 13800
aatctgaaaa aataaaagac ggtactcacc aaggccatgg cgaccagcgt gcgttcttct 13860
ctgttgtttg tagtagtatg atgaggcgcg tgtaccgga gggtcctcct ccctcgtacg 13920
agagcgtgat gcagcaggcg gtggcggcgg cgatgcagcc cccgctggag gcgccttacg 13980
tgccccgcg gtacctggcg cctacggagg ggcggaacag cattcggttac tcggagctgg 14040
cacccttgta cgataccacc cggttgtacc tgggtggaaa caagtcggca gacatcgctt 14100
cgctgaacta ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt 14160

tcacccccac ggaggccagc acccagacca tcaactttga cgagcgctcg cgggtggggcg 14220
 gccagctgaa aaccatcatg cacaccaaca tgcccaacgt gaacgagttc atgtacagca 14280
 acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc caacgggggtg gatgatgatt 14340
 atgatggtag tcaggacgag ctgacctacg agtgggtgga gtttgagctg cccgagggca 14400
 acttctcggg gaccatgacc atcgatctga tgaacaacgc catcatcgac aactacttgg 14460
 cgggtggggcg gcagaacggg gtgctggaga gcgacatcgg cgtgaagttc gacacgcgca 14520
 acttccggct gggctgggac cccgtgaccg agctgggtgat gccgggctg tacaccaacg 14580
 aggccttcca ccccagacatc gtcctgctgc ccggctgcgg cgtggacttc accgagagcc 14640
 gcctcagcaa cctgctggg atccgcaagc ggcagccctt ccaggagggc ttccagatcc 14700
 tgtacgagga cctggagggg ggcaacatcc ccgcgctctt ggatgtcgaa gcctacgaga 14760
 aaagcaagga ggatagcacc gccgcggcga ccgcagccgt ggccaccgcc tctaccgagg 14820
 tgcggggcga taattttgct agcgctgcgg cagcggccga ggcggctgaa accgaaagta 14880
 agatagtcac ccagccgggtg gagaaggaca gcaaggacag gagctacaac gtgctcgcgg 14940
 acaagaaaaa caccgcctac cgcagctggg acctggccta caactacggc gaccccagaga 15000
 agggcgtgcg ctcttgagc ctgctacca cctcggacgt cacctgcggc gtggagcaag 15060
 tctactggtc gctgcccagc atgatgcaag acccgggtcac cttccgctcc acgctcaag 15120
 ttagcaacta cccgggtggg ggcgccgagc tcctgcccgt ctactccaag agcttcttca 15180
 acgagcaggc cgtctactcg cagcagctgc gcgccttcac ctgctcacg cacgtcttca 15240
 accgcttccc cgagaaccag atcctcgtcc gcccgcccgc gccaccatt accaccgtca 15300
 gtgaaaacgt tcctgctctc acagatcacg ggaccctgcc gctgcgcagc agtatccggg 15360
 gagtccagcg cgtgaccgtc actgacgcca gacgccgcac ctgcccctac gtctacaagg 15420
 ccctgggctg agtcgcgccg cgcgtcctct cgagccgcac cttctaaaaa atgtccattc 15480
 tcatctcgcc cagtaataac accggttggg gcctgcgcgc gccagcaag atgtacggag 15540
 gcgctcgcca acgctccacg caacaccccg tgcgcgtgcg cgggcacttc cgcgtccct 15600
 ggggcgccct caagggccgc gtgcgctcgc gcaccaccgt cgacgacgtg atcgaccagg 15660
 tgggtggccga cgcgcgcaac tacacgcccg ccgccgcgcc cgtctccacc gtggacgccg 15720
 tcatcgacag cgtggtggcc gacgcgcgcc ggtacgcccg caccaagagc cggcggcgcc 15780
 gcatcgcccg gcggcaccgg agcacccccg ccatgcgcgc ggcgcgagcc ttgctgcgca 15840
 gggccaggcg cacgggacgc agggccatgc tcagggcggc cagacgcgcg gcctccggca 15900
 gcagcagcgc cggcaggacc cgcagacgcg cggccacggc ggcggcggcg gccatcgcca 15960
 gcatgtcccg cccgcggcgc ggcaacgtgt actgggtgcg cgacgccgcc accggtgtgc 16020
 gcgtgcccgt gcgcacccgc cccctcgca cttgaagatg ctgacttcgc gatgttgatg 16080
 tgtcccagcg gcgaggagga tgtccaagcg caaatacaag gaagagatgc tccagggtcat 16140
 cgcgcctgag atctacggcc ccgcggcggc ggtgaaggag gaaagaaagc cccgcaaact 16200
 gaagcgggtc aaaaaggaca aaaaggagga ggaagatgac ggactggtgg agtttgtgcg 16260

cgagttcgcc ccccggcggc gcgtgcagtg gcgcgggcgg aaagtgaac cgggtgctgcg 16320
gcccggcacc acggtggtct tcacgcccgg cgagcggtcc ggctccgcct ccaagcgctc 16380
ctacgacgag gtgtacgggg acgaggacat cctcgagcag gcggtcgagc gtctgggcga 16440
gtttgcgtac ggcaagcgca gccgccccgc gcccttgaaa gaggaggcgg tgtccatccc 16500
gctggaccac ggcaacccca cgccgagcct gaagccgggtg accctgcagc aggtgctacc 16560
gagcgcggcg ccgcgccggg gcttcaagcg cgagggcggc gaggatctgt acccgaccat 16620
gcagctgatg gtgcccagc gccagaagct ggaggacgtg ctggagcaca tgaagggtgga 16680
ccccgaggtg cagccccgag tcaagggtgcg gcccatcaag cagggtggccc cgggcctggg 16740
cgtgcagacc gtggacatca agatccccac ggagcccatg gaaacgcaga ccgagcccgt 16800
gaagcccagc accagcacca tggagggtgca gacggatccc tggatgccag caccagcttc 16860
caccagcact cgccgaagac gcaagtacgg cgcgccagc ctgctgatgc ccaactacgc 16920
gctgcatcct tccatcatcc ccacgccggg ctaccgcggc acgcgcttct accgcggcta 16980
caccagcagc cgccgccgca agaccaccac ccgccgccgt cgtcgagacc gccgcagcag 17040
caccgcgact tccgccttgg tgcggagagt gtatcgagc gggcgcgagc ctctgaccct 17100
gccgcgcg cgctaccacc cgagcatcgc catttaacta ccgcctccta cttgcagata 17160
tggccctcac atgccgcctc cgcggtcccca ttacgggcta ccgaggaaga aagccgcgcc 17220
gtagaaggct gacggggaac gggctgcgtc gccatcacca ccggcggcgg cgcgccatca 17280
gcaagcgggt ggggggaggc ttcctgcccg cgctgatccc catcatcgcc gcggcgatcg 17340
gggcgatccc cgcatagct tccgtggcgg tgcaggcctc tcagcgccac tgagacacaa 17400
aaaagcatgg atttgtaata aaaaaaaaaa tggactgacg ctcttggtcc tgtgatgtgt 17460
gttttttagat ggaagacatc aatttttcgt ccctggcacc gcgacacggc acgcggccgt 17520
ttatgggcac ctggagcgac atcggcaaca gccaaactgaa cgggggcgcc ttcaattgga 17580
gcagtctctg gagcgggctt aagaatttcg ggtccacgct caaacctat ggcaacaagg 17640
cgtggaacag cagcacaggc caggcgctga gggaaaagct gaaagaacag aacttcagc 17700
agaagggtgt tgatggcctg gcctcaggca tcaacggggt ggttgacctg gccaaaccagg 17760
ccgtgcagaa acagatcaac agccgcctgg acgcggtccc gcccgcgggg tccgtggaga 17820
tgccccaggt ggaggaggag ctgcctcccc tggacaagcg cggcgacaag cgaccgcgtc 17880
ccgacgcgga ggagacgctg ctgacgcaca cggacgagcc gccccgtac gaggaggcgg 17940
tgaaactggg cctgcccacc acgcggcccc tggcgctctt ggccaccgga gtgctgaaac 18000
ccagcagcag ccagcccgcg accctggact tgcctccgcc tcgcccctcc acagtggcta 18060
agcccctgcc gccggtggcc gtcgcgtcgc gcgcccccg aggcgcgccc caggcgaact 18120
ggcagagcac tctgaacagc atcgtgggtc tgggagtgca gagtgtgaag cgccgccgct 18180
gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt atgtccgccg 18240
accagaagga ggagtgtgaa gaggcgcgtc gccgagttgc aagatggcca ccccatcgat 18300

gctgccccag tgggcggtaca tgcacatcgc cggacaggac gcttcggagt acctgagtcc 18360
gggtctggtg cagttcgccc gcgccacaga cacctacttc agtctgggga acaagtttag 18420
gaacccccacg gtggcgccca cgcacgatgt gaccaccgac cgcagccagc ggctgacgct 18480
gcgcttcgtg cccgtggacc gcgaggacaa cacctactcg taaaaagtgc gctacacgct 18540
ggccgtgggc gacaaccgcg tgctggacat ggccagcacc tactttgaca tccgcggcgt 18600
gctggaccgg ggccctagct tcaaaccccta ctctggcacc gcctacaaca gcctagctcc 18660
caaggaggct cccaattcca gccagtggga gcaagcaaaa acaggcaatg ggggaactat 18720
ggaaacacac acatatggtg tggccccaat gggcggagag aatattacaa aagatggtct 18780
tcaaattgga actgacgtta cagcgaatca gaataaacca atttatgccg aaaaaacatt 18840
tcaaccagaa ccgcaagtag gagaagaaaa ttggcaagaa actgaaaact tttatggcgg 18900
tagagctctt aaaaaagaca caaacatgaa accttgctat ggctcctatg ctagaccac 18960
caatgaaaaa ggaggtcaag ctaaaactta agttggagat gatggagttc caaccaaga 19020
attcgacata gacctggctt tctttgatac tcccgggtggc accgtgaacg gtcaagacga 19080
gtataaagca gacattgtca tgtataccga aaacacgtat ttggaaactc cagacacgca 19140
tgtggtatac aaaccaggca aggatgatgc aagttctgaa attaacctgg ttcagcagtc 19200
tatgcccac agacccaact acattgggtt cagggaacaac tttatcggtc ttatgtacta 19260
caacagcact ggcaatatgg gtgtgcttgc tggtcaggcc tcccagctga atgctgtggt 19320
tgatttgcaa gacagaaaca ccgagctgtc ctaccagctc ttgcttgact ctttgggtga 19380
cagaacccgg tatttcagta tgtggaacca ggcgggtggac agttatgacc ccgatgtgcg 19440
catcatcgaa aaccatggtg tggaggatga attgccaaac tattgcttcc ccttggacgg 19500
ctctggcact aacgccgcat accaagggtg gaaagtaaaa gatggtcaag atggtgatgt 19560
tgagagtga tgggaaaatg acgatactgt tgcagctcga aatcaattat gtaaaggtaa 19620
cattttcgcc atggagatta atctccaggc taacctgtgg agaagtttcc tctactcgaa 19680
cgtggccctg tacctgcccg actcctacaa gtacacgccg accaacgtca cgctgccgac 19740
caacaccaac acctacgatt acatgaatgg cagagtgaac cctccctcgc tggtagacgc 19800
ctacctcaac atcggggcgc gctgggtcgt ggaccccatg gacaacgtca accccttcaa 19860
ccaccaccgc aacgcgggccc tgcgtaccg ctccatgctc ctgggcaacg ggcgctacgt 19920
gcccttccac atccagggtg cccaaaagtt tttcgccatc aagagcctcc tgctcctgcc 19980
cgggtcctac acctacgagt ggaacttccg caaggacgtc aacatgatcc tgcagagctc 20040
cctaggcaac gacctgcgca cggacggggc ctccatcgcc ttcaccagca tcaacctcta 20100
cgccaccttc ttcccatg cgcacaacac cgcctccacg ctcgaggcca tgctgcgcaa 20160
cgacaccaac gaccagtcct tcaacgacta cctctcggcg gccaacatgc tctaccccat 20220
cccggccaac gccaccaacg tgcccatctc catcccctcg cgcaactggg ccgccttccg 20280
cggatggtcc ttcacgcgcc tgaagaccg cgagacgccc tcgctcggct ccgggttcga 20340
cccctacttc gtctactcgg gctccatccc ctacctagac ggcaccttct acctcaacca 20400

caccttcaag aaggtctcca tcaccttcga ctctccgctc agctggcccc gcaacgaccg 20460
cctcctgacg cccaacgagt tcgaaatcaa gcgcaccgctc gacggagagg gatacaacgt 20520
ggcccagtg c aacatgacca aggactgggtt cctgggtccag atgctggccc actacaacat 20580
cggctaccag ggcttctacg tgcccagagg ctacaaggac cgcattgtact ccttcttccg 20640
caacttccag cccatgagcc gccaggctcgt ggacgaggtc aactacaagg actaccaggc 20700
cgtcaccttg gcctaccagc acaacaactc gggcttcgctc ggctacctcg cgcccaccat 20760
gcgccagggc cagccctacc ccgccaacta cccctacccg ctcatcggca agagcgccgt 20820
cgccagcgtc acccagaaaa agttcctctg cgaccgggtc atgtggcgca tcccccttc 20880
cagcaacttc atgtccatgg gcgcgctcac cgacctcggc cagaacatgc tctacgcaa 20940
ctccgcccac gcgctagaca tgaatttcga agtcgacccc atggatgagt ccacccttct 21000
ctatgttgctc ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc gcggcgtcat 21060
cgaagccgtc tacctgcgca cgcccttctc ggccggcaac gccaccacct aagccgctct 21120
tgcttcttgc aagatgacgg cgggctccgg cgagcaggag ctcagggcca tcctccgca 21180
cctgggctgc gggccctgct tcctgggcac cttcgacaag cgcttccctg gattcatggc 21240
cccgacaaag ctggcctgcg ccatcgtgaa cacggccggc cgcgagaccg ggggcgagca 21300
ctggctggcc ttcgcctgga acccgcgctc ccacacatgc tacctcttcg accccttcgg 21360
gttctcggac gagcgctca agcagatcta ccagttcgag tacgagggcc tgctgcgtcg 21420
cagcgccctg gccaccgagg accgctgcgt caccctggaa aagtccacc agaccgtgca 21480
gggtccgcgc tcggccgcct gcgggctctt ctgctgcatg ttcctgcacg ccttcgtgca 21540
ctggcccagc cgcccatgg acaagaacc caccatgaac ttactgacgg gggtgcccaa 21600
cggcatgctc cagtcgcccc aggtggaacc caccctgcgc cgcaaccagg aagcgctcta 21660
ccgcttctc aatgccact ccgcctactt tcgctccac cgcgcgcgca tcgagaaggc 21720
caccgccttc gaccgatga atcaagacat gtaaaaaacc ggtgtgtgta tgtgaatgct 21780
ttattcataa taaacagcac atgtttatgc cacttctctc gaggtctga ctttatttag 21840
aaatcgaagg ggttctgccc gctctcggca tggcccgcgg gcagggatac gttgcggaac 21900
tggtacttgg gcagccactt gaactcgggg atcagcagct tgggcacggg gaggtcgggg 21960
aacgagtcgc tccacagctt gcgcgtgagt tgcagggcgc ccagcaggtc gggcgcgagg 22020
atcttgaaat cgcagttggg acccgcttc tcgcgcgag agttgcggta cacgggggtg 22080
cagcactgga acaccatcag ggccgggtgc ttcacgcttg ccagcaccgt cgcgtcggtg 22140
atgccctcca cgtccagatc ctccgcttg gccatcccga agggggctcat cttgcaggtc 22200
tgccgcccc tgctgggcac gcagccgggc ttgtggttg aatcgagtg cagggggatc 22260
agcatcatct gggcctgctc ggagctcatg cccgggtaca tggccttcat gaaagcctcc 22320
agctggcgga aggcctgctg cgccttgccg ccctcgggtga agaagacccc gcaggacttg 22380
ctagagaact ggttggtggc gcagccggcg tcgtgcacgc agcagcgcg gtcgttggtg 22440

gccagctgca ccacgctgcg cccccagcgg ttctgggtga tcttggtccc gttgggggttc 22500
tccttcagcg cgcgctgccc gttctcgctc gccacatcca tctcgatagt gtgctccttc 22560
tggtatcatca cgggtcccggt caggcaccgc agcttgccct cggcttcggt gcagccgtgc 22620
agccacagcg cgcagccggt gactcccgag ttcttggtgg cgatctggga gtgcgagtgc 22680
acgaagccct gcaggaagcg gcccatcatc gcggtcaggg tcttggtgct ggtgaaggtc 22740
agcgggatgc cgcggtgctc ctggttcaca tacaggtggc agatgcggcg gtacacctcg 22800
ccctgctcgg gcatcagctg gaaggcggac ttcaggtcgc tctccacgcg gtaccgggtcc 22860
atcagcagcg tcatcacttc catgcccttc tcccaggccg aaacgatcgg caggctcagg 22920
gggttcttca ccgccattgt catcttagtc gccgccgccg aggtcagggg gtcgttctcg 22980
tccaggtctt caaacactcg cttgccgtcc ttctcgatga tgcgcacggg gggaaagctg 23040
aagcccacgg ccgccagctc ctctcgggc tgcctttcgt cctcgctgtc ctggctgatg 23100
tcttgcaaag gcacatgctt ggtcttgcg ggtttctttt tgggcggcag aggcggcggc 23160
gatgtgctgg gagagcgcg gttctcgttc accacgacta tttcttcttc ttggccgtcg 23220
tccgagacca cgcggcggtg ggcagagccc cttccgcgtt cgggggtgcg ctctggcg 23340
cgctgctctg actgacttcc tccgcggccg gccattgtgt tctcctaggg agcaacaaca 23400
agcatggaga ctgagccatc gtcgccaaca tcgccatctg ccccgccgc caccgccgac 23460
gagaaccagc agcagaatga aagcttaacc gccccgccgc ccagccccac ctccgacgcc 23520
gcggccccag acatgcaaga gatggaggaa tccatcgaga ttgacctggg ctacgtgacg 23580
cccgcggagc acgaggagga gctggcagcg cgcttttcag ccccggaaga gaaccaccaa 23640
gagcagccag agcaggaagc agagaacgag cagaaccagg ctgggcacga gcatggcgac 23700
tacctgagcg gggcagagga cgtgctcatc aagcatctgg ccgccaatg catcatcgtc 23760
aaggacgcgc tgctcgaccg cgccgaggtg cccctcagcg tggcggagct cagccgcgcc 23820
tacgagcgca acctcttctc gccgcgcgtg cccccaagc gccagcccaa cggcacctgt 23880
gagcccaacc cgcgcctcaa cttctacccg gtcttcgcgg tgcccagggc cctggccacc 23940
taccacctct ttttcaagaa ccaaaggatc cccgtctcct gccgcgcaa ccgcacccgc 24000
gccgacgcc tgctcaacct gggccccggc gccgcctac ctgatatcac ctcttgga 24060
gaggttccca agatcttcga gggcttgggc agcgacgaga ctcgggccgc gaacgctctg 24120
caaggaagcg gagaggagca tgagcaccac agcgccttg tggagtggga aggcgacaac 24180
gcgcgcctgg cggctctcaa gcgcacggtc gagctgacct acttcgccta cccggcgctc 24240
aacctgcccc ccaaggtcat gagcgccgtc atggaccagg tgctcatcaa gcgcgcctcg 24300
cccctctcgg aggaggagat gcaggacccc gagagttcgg acgagggcaa gcccggtggtc 24360
agcgacgagc agctggcgcg ctggctggga gcgagtagca ccccccagag cctggaagag 24420
cggcgcaagc tcatgatggc cgtggtcctg gtgaccgtgg agctggagtg tctgcgccgc 24480
ttctttgccg acgcggagac cctgcgcaag gtcgaggaga acctgcacta cctcttcagg 24540

cacgggttcg tgcgccaggc ctgcaagatc tccaacgtgg agctgaccaa cctggtctcc 24600
 tacatgggca tcctgcacga gaaccgcctg gggcaaaacg tgctgcacac caccctgcgc 24660
 ggggaggccc gccgcgacta catccgcgac tgcgtctacc tgtacctctg ccacacctgg 24720
 cagacgggca tgggcgtgtg gcagcagtgc ctggaggagc agaacctgaa agagctctgc 24780
 aagctcctgc agaagaacct caaggccctg tggaccgggt tcgacgagcg taccaccgcc 24840
 tcggacctgg ccgacctcat cttccccgag cgcttgcggc tgacgctgcg caacgggctg 24900
 cccgacttta tgagccaaag catgttgcaa aactttcgct ctttcatcct cgaacgctcc 24960
 gggatcctgc ccgccacctg ctccgcgctg ccctcggact tcgtgccgct gaccttccgc 25020
 gagtgtcccc cgccgctctg gagccactgc tacttgctgc gcctggccaa ctacctggcc 25080
 taccactcgg acgtgatcga ggacgtcagc ggcgagggtc tgctggagtg cactgcccgc 25140
 tgcaacctct gcacgccga ccgctccctg gcctgcaacc cccagctgct gagcgagacc 25200
 cagatcatcg gcaccttcga gttgcaaggc cccggcgacg gcgagggcaa ggggggtctg 25260
 aaactcacc cggggctgtg gacctcggcc tacttgcgca agttcgtgcc cgaggactac 25320
 catcccttcg agatcaggtt ctacgaggac caatcccagc cgccaaggc cgagctgtcg 25380
 gcctgcgtca tcaccaggg ggccatcctg gcccaattgc aagccatcca gaaatcccgc 25440
 caagaatttc tgctgaaaaa gggccacggg gtctacttgg acccccagac cggagaggag 25500
 ctcaaccca gcttcccca ggatgccccg aggaagcagc aagaagctga aagtggagct 25560
 gccgccgccg gaggatttgg aggaagactg ggagagcagt caggcagagg aggaggagat 25620
 ggaagactgg gacagcactc aggcagagga ggacagcctg caagacagtc tggaggagga 25680
 agacgaggtg gaggaggcag aggaagaagc agccgccgcc agaccgtcgt cctcggcgga 25740
 gaaagcaagc agcacggata ccatctccgc tccgggtcgg ggtcgcggcg gccgggcca 25800
 cagtaggtgg gacgagaccg ggcgcttccc gaacccacc acccagaccg gtaagaagga 25860
 gcggcagggg tacaagtcct ggcgggggca caaaaacgcc atcgtctcct gcttgcaagc 25920
 ctgcgggggc aacatctcct tcaccggcg ctacctgtc ttccaccgcg ggggtgaactt 25980
 cccccgaac atcttgcat actaccgtca cctccacagc ccctactact gtttccaaga 26040
 agaggcagaa acccagcagc agcagaaaac cagcggcagc agcagctaga aaatccacag 26100
 cggcggcagg tggactgagg atcgcggcga acgagccggc gcagaccgcg gagctgagga 26160
 accggatctt tcccaccctc tatgccatct tccagcagag tcgggggcag gagcaggaac 26220
 tgaaagtcaa gaaccgttct ctgcgctcgc tcaccgcag ttgtctgtat cacaagagcg 26280
 aagaccaact tcagcgact ctgaggagc cggaggctct cttcaacaag tactgcgcgc 26340
 tcactcttaa agagtagccc gcgcccgc acacacggaa aaaggcggga attacgtcac 26400
 cacctgcgcc cttcgcccga ccatcatgag caaagagatt cccacgcctt acatgtggag 26460
 ctaccagccc cagatgggccc tggccgcccg cggcggccag gactactcca cccgcatgaa 26520
 ctggctcagt gccgggcccg cgatgatctc acgggtgaat gacatccgcg cccaccgaaa 26580

ccagatactc ctagaacagt cagcgatcac cgccacgccc cgccatcacc ttaatccgcg 26640
 taattggccc gccgccctgg tgtaccagga aattccccag cccacgaccg tactacttcc 26700
 gcgagacgcc caggccgaag tccagctgac taactcaggt gtccagctgg ccggcggcgc 26760
 cgccctgtgt cgtcaccgcc ccgctcaggg tataaagcgg ctgggtgatcc gaggcagagg 26820
 cacacagctc aacgacgagg tgggtgagctc ttcgctgggt ctgcgacctg acggagtctt 26880
 ccaactcgcc ggatcgggga gatcttcctt cacgcctcgt caggccgtcc tgactttgga 26940
 gagttcgtcc tcgcagcccc gtcgggcggt catcggcact ctccagttcg tggaggagtt 27000
 cactccctcg gtctacttca accccttctc cggctcccc ggccactacc cggacgagtt 27060
 catcccgaac ttcgacgcca tcagcgagtc ggtggacggc tacgattgaa tgtcccatgg 27120
 tggcgagct gacctagctc ggcttcgaca cctggaccac tgccgccgct tccgctgctt 27180
 cgctcgggat ctgcgccagt ttgcctactt tgagctgccc gaggagcacc ctcagggccc 27240
 agcccacgga gtgcggatca tcgtcgaagg gggcctcgac tcccacctgc ttcggatctt 27300
 cagccagcga ccgatcctgg tcgagcgca acaaggacag acccttctta ctttgtactg 27360
 catctgcaac caccgccgcc tgcataaaag tctttgttgt ctgctgtgta ctgagtataa 27420
 taaaagctga gatcagcgac tactccggac tcgattgtgg tgttcctgct atcaaccggt 27480
 ccctgttctt caccgggaac gagaccgagc tccagctcca gtgtaagccc cacaagaagt 27540
 acctcacctg gctgttccag ggctccccga tcgccgttgt caaccactgc gacaacgacg 27600
 gagtcctgct gagcggccct gccaacctta ctttttccac ccgcagaagc aagctccagc 27660
 tcttccaacc cttcctcccc gggacctatc agtgcgcttc aggaccctgc catcacacct 27720
 tccacctgat cccgaatacc acagcgccgc tccccgctac taacaaccaa actaccacc 27780
 aacgccaccg tcgcgacctt tcctctgaat ctaataccac taccggaggt gagctccgag 27840
 gtcgaccaac ctctgggatt tactacggcc cctgggaggt ggtgggggta atagcgctag 27900
 gcctagttgc ggggtgggctt ttggttctct gctacctata cctcccttgc tgttcgtact 27960
 tagtggtgct gtgttgctgg ttttaagaaat ggggaagatc accctagtga gctgcggtgc 28020
 gctggtggcg gtgttgcttt cgattgtggg actgggcggc gcggctgtag tgaaggagaa 28080
 ggccgatccc tgcttgcat tcaatcccaa caaatgccag ctgagttttc agcccgatgg 28140
 caatcgggtgc gcggtactga tcaagtgcgg atgggaatgc gagaacgtga gaatcgagta 28200
 caataacaag actcggaaac atactctcgc gtccgtgtgg cagcccgggg accccgagtg 28260
 gtacaccgtc tctgtccccg gtgctgacgg ctccccgcgc accgtgaata atactttcat 28320
 ttttgcgcac atgtgcaaca cggatcatgt gatgagcaag cagtacgata tgtggcccc 28380
 cacgaaggag aacatcgtgg tcttctccat cgcttacagc ctgtgcacgg cgctaatac 28440
 cgctatcgtg tgctgagca ttcacatgct catcgtatt cgccccagaa ataatgccga 28500
 gaaagagaaa cagccataac acgttttttc acacaccttg tttttacaga caatgcgtct 28560
 gttaaatttt ttaaacattg tgctcagtat tgcttatgcc tctggttatg caaacataca 28620
 gaaaaccctt tatgtaggat ctgatggtac actagagggt acccaatcac aagccaaggt 28680

tgcattggtat ttttatagaa ccaacactga tccagttaaa ctttgtaagg gtgaattgcc 28740
 gcgtacacat aaaactccac ttacatttag ttgcagcaat aataatctta cacttttttc 28800
 aattacaaaa caatatactg gtacttatta cagtacaaac tttcatacag gacaagataa 28860
 atattatact gttaaggtag aaaatcctac cactcctaga actaccacca ccaccactac 28920
 tgcaaagccc actgtgaaaa ctacaactag gaccaccaca actacagaaa ccaccaccag 28980
 cacaacactt gctgcaacta cacacacaca cactaagcta accttacaga ccactaatga 29040
 tttgatcgcc ctgctgcaaa agggggataa cagcaccact tccaatgagg agatacccaa 29100
 atccatgatt ggcattattg ttgctgtagt ggtgtgcatg ttgatcatcg ccttggtgcat 29160
 ggtgtactat gccttctgct acagaaagca cagactgaac gacaagctgg aacacttact 29220
 aagtgttgaa ttttaatttt ttagaacat gaagatccta ggccttttta gtttttctat 29280
 cattacctct gctctttgtg aatcagtga tagagatgtt actattacca ctggttctaa 29340
 ttatacactg aaagggccac cctcaggtat gctttcgtgg tattgctatt ttggaactga 29400
 cactgatcaa actgaattat gcaattttca aaaaggcaaa acctcaaact ctaaaatctc 29460
 taattatcaa tgcaatggca ctgatctgat actactcaat gtcacgaaag catatggtgg 29520
 cagttattat tgccctggac aaaacactga agaaatgatt ttttacaaag tggaagtgg 29580
 tgatcccact acaccacca ccaccacaac tattcatacc acacacacag aacaaacacc 29640
 agaggcaaca gaagcagagt tggccttcca gggtcacgga gattcctttg ctgtcaatac 29700
 ccctacacc gatcagcggg gtccggggcc gctagtcagc ggcattgtcg gtgtgctttc 29760
 gggattagca gtcataatca tctgcatgtt cttttttgct tgctgctata gaaggcttta 29820
 ccgacaaaaa tcagaccac tgctgaacct ctatgtttta ttttttccag agccatgaag 29880
 gcagttagcg ctctagtttt ttgttctttg attggcattg tttttaatag taaaattacc 29940
 agagttagct ttattaaaca tgtaaatgta actgaaggag ataacatcac actagcaggt 30000
 gtagaagggt ctcaaaacac cacctggaca aaataccatc taggatggag agatatttgc 30060
 acctggaatg taacttatta ttgcatagga gttaatctta ccattgttaa cgctaacc 30120
 tctcagaatg ggttaattaa aggacagagt gttagtgtga ccagtgatgg gtactatacc 30180
 cagcatagtt ttaactacaa cattactgtc ataccactgc ctacgcctag cccacctagc 30240
 actaccacac agacaaccac atacagtaca tcaaatcagc ctaccaccac tacagcagca 30300
 gaggttgcca gctcgtctgg ggtccgagtg gcatttttga tgttggcccc atctagcagt 30360
 cccactgcta gtaccaatga gcagactact gaatttttgt cactgtcga gagccacacc 30420
 acagctacct ccagtgcctt ctctagcacc gccaatctct cctcgctttc ctctacacca 30480
 atcagccccg ctactactcc tagccccgt cctcttccca ctcccctgaa gcaaacagac 30540
 ggcggcatgc aatggcagat caccctgctc attgtgatcg ggttggtcat cctggccgtg 30600
 ttgctctact acatcttctg ccgccgcatt cccaacgcgc accgcaagcc ggcctacaag 30660
 cccatcgta tcgggcagcc ggagccgctt caggtggaag ggggtctaag gaatcttctc 30720

ttctcttttta cagtatggtg attgaactat gattcctaga caattcttga tcactattct 30780
 tatctgcctc ctccaagtct gtgccaccct cgctctggtg gccaacgcca gtccagactg 30840
 tattgggccc ttgcctcct acgtgctctt tgccttcgtc acctgcatct gctgctgtag 30900
 catagtctgc ctgcttatca cttctttcca gttcattgac tggatctttg tgcgcatcgc 30960
 ctacctgcgc caccaccccc agtaccgcga ccagcgagtg gcgcagctgc tcaggctcct 31020
 ctgataagca tgcgggctct gctacttctc gcgcttctgc tgttagtgct cccccgtccc 31080
 gtcgaccccc ggtccccac tcagtcccc gaggaggttc gcaaatacaa attccaagaa 31140
 ccctggaaat tcctcaaatg ctaccgcaa aaatcagaca tgcattcccag ctggatcatg 31200
 atcattggga tcgtgaacat tctggcctgc accctcatct cttttgtgat ttaccctgc 31260
 tttgactttg gttggaactc gccagaggcg ctctatctcc cgctgaacc tgacacacca 31320
 ccacagcagc aacctcaggc acacgacta ccaccaccac agcctaggcc acaatacatg 31380
 cccatattag actatgaggc cgagccacag cgacctatgc tccccgctat tagttacttc 31440
 aatctaaccg gcggagatga ctgacctact ggccaataac aacgtcaacg accttctcct 31500
 ggacatggac ggccgcgcct cggagcagcg actcgcccaa cttcgcattc gtcagcagca 31560
 ggagagagcc gtcaaggagc tgcaggacgg catagccatc caccagtga agagaggcat 31620
 cttctgcctg gtgaaacagg ccaagatctc ctacgaggtc acccagaccg accatcgcct 31680
 ctctacgag ctctgcagc agcgccagaa gttcacctgc ctggtcggag tcaaccccat 31740
 cgtcatcacc cagcagtcgg gcgatacaa ggggtgcatc cactgctcct gcgactcccc 31800
 cgactgcgtc cacactctga tcaagacct ctgcggcctc cgcgacctc tccccatgaa 31860
 ctaatcacc cttatccag tgaaataaag atcatattga tgatgattta aataaaaaaa 31920
 ataatcattt gatttgaaat aaagatacaa tcatattgat gatttgagtt taacaaaaat 31980
 aaagaatcac ttacttgaat tctgatacca ggtctctgtc catgttttct gccaacacca 32040
 cctcactccc ctcttcccag ctctggtact gcaggccccg gcgggctgca aacttcctcc 32100
 acacgctgaa ggggatgtca aattcctcct gtccctcaat cttcatttta tcttctatca 32160
 gatgtccaaa aagcgcgtcc gggtagtgat tgacttcgac cccgtctacc cctacgatgc 32220
 agacaacgca ccgaccgtgc cttcatcaa ccccccttc gtctcttcag atggattcca 32280
 agagaagccc ctgggggtgt tgtccctgcg actggctgac cccgtcacca ccaagaacgg 32340
 ggaaatcacc ctcaagctgg gagaggggtt ggacctcgac tcgtcgggaa aactcatctc 32400
 caacacggcc accaaggccg ccgcccctct cagtatttca aacaacacca tttcccttaa 32460
 aactgctgcc ctttctaca acaacaatgg aactttaagc ctcaatgtct ccacaccatt 32520
 agcagtattt cccacattta acactttagg cataagtctt ggaaacggtc ttcagacttc 32580
 aaataagttg ttgactgtac aactaactca tcctcttaca ttcagctcaa atagcatcac 32640
 agtaaaaaca gacaaagggc tatatatata ctccagtggg aacagaggac ttgaggctaa 32700
 tataagccta aaaagaggac tagtttttga cggtaatgct attgcaacat atattggaaa 32760
 tggcttagac tatggatctt atgatagtga tggaaaaaca agaccgtaa ttacaaaaat 32820

tggagcagga ttaaattttg atgctaacaa agcaatagct gtcaaactag gcacagggttt 32880
 aagttttgac tccgctggtg ccttgacagc tggaaacaaa caggatgaca agctaacact 32940
 ttggactacc cctgacccaa gccctaattg tcaattactt tcagacagag atgccaaatt 33000
 tactctctgt cttacaaaat gcggtagtca aatactaggc actgtggcag tggcggctgt 33060
 tactgtagga tcagcactaa atccaattaa tgacacagtc aaaagcgcca tagttttcct 33120
 tagatttgat tccgatggtg tactcatgtc aaactcatca atggtaggtg attactggaa 33180
 ctttagggag ggacagacca ctcaaagtgt agcctataca aatgctgtgg gattcatgcc 33240
 aaatataggt gcatatccaa aaacccaaag taaaacacct aaaaatagca tagtcagtca 33300
 ggtatattta actggagaaa ctactatgcc aatgacacta accataactt tcaatggcac 33360
 tgatgaaaaa gacacaaccc cagttagcac ctactctatg acttttacat ggcatggac 33420
 tggagactat aaggacaaaa atattacctt tgctaccaac tcattctctt tttcctacat 33480
 cgcccaggaa taatcccacc cagcaagcca accccttttc ccaccacctt tgtctatatg 33540
 gaaactctga aacagaaaaa taaagttcaa gtgttttatt gaatcaacag ttttacagga 33600
 ctcgagcagt tatttttcct ccaccctccc aggacatgga atacaccacc ctctcccccc 33660
 gcacagcctt gaacatctga atgccattgg tgatggacat gcttttggtc tccacgttcc 33720
 acacagtttc agagcgagcc agtctcggat cggtcaggga gatgaaaccc tccgggcaact 33780
 cccgcatctg cacctcacag ctcaacagct gaggattgtc ctcggtggc gggatcacgg 33840
 ttatctggaa gaagcagaag agcggcgggtg ggaatcatag tccgcgaacg ggatcggccg 33900
 gtggtgtcgc atcaggcccc gcagcagtcg ctgccgccgc cgctccgtca agctgctgct 33960
 caggggggttc ggggtccaggg actccctcag catgatgcc acggccctca gcatcagtcg 34020
 tctggtgcgg cgggcgcagc agcgcagtcg aatctcgtc aggtcactgc agtacgtgca 34080
 acacaggacc accaggttgt tcaacagtcc atagttcaac acgctccagc cgaaactcat 34140
 cgcggaagg atgctacca cgtggccgtc gtaccagatc ctcaggtaaa tcaagtggcg 34200
 ctccctccag aagacgctgc ccatgtacat gatctcctt ggcatgtggc ggttcaccac 34260
 ctcccgttac cacatcacc tctggttgaa catgcagccc cggatgatcc tgcggaacca 34320
 cagggccagc accgccccgc ccgcatgca gcgaagagac cccggatccc ggcaatgaca 34380
 atggaggacc caccgctcgt acccggtgat catctgggag ctgaacaagt ctatgttggc 34440
 acagcacagg catatgctca tgcattctt cagcactctc agctcctcgg ggggtcaaac 34500
 catatcccag ggcacgggga actcttgag gacagcgaac cccgcagaac agggcaatcc 34560
 tcgcacataa cttacattgt gcatggacag ggtatcgcaa tcaggcagca ccgggtgatc 34620
 ctccaccaga gaagcgcggg tctcgggtctc ctacagcgt ggtaaggggg ccggccgata 34680
 cgggtgatgg cgggacgcgg ctgatcgtgt tctcgaccgt gtcatgatgc agttgctttc 34740
 ggacattttc gtacttgctg tagcagaacc tgggtccggc gctgcacacc gatcgccggc 34800
 ggcggtctcg gcgcttgaa cgctcgggtgt taaagttgta aaacagccac tctctcagac 34860

cggtgcagcag atctaggggcc tcaggagtga tgaagatccc atcatgcctg atagctctga 34920
 tcacatcgac caccgtggaa tgggccaggc ccagccagat gatgcaattt tgttgggttt 34980
 cgggtgacggc gggggaggga agaacaggaa gaaccatgat taacttttaa tccaaacggg 35040
 ctctggagcac ttcaaaatga aggtcacgga gatggcacct ctgcccccg ctgtgttggg 35100
 ggaaaataac agccagggtca aagggtgatac ggttctcgag atgttccacg gtggcttcca 35160
 gcaaagcctc cacgcgcaca tccagaaaca agacaatagc gaaagcgga gggttctcta 35220
 attcctcaac catcatgtta cactcctgca ccatccccag ataattttca tttttccagc 35280
 cttgaatgat tcgaactagt tcctgaggta aatccaagcc agccatgata aaaagctcgc 35340
 gcagagcacc ctccaccggc attcttaagc acaccctcat aattccaaga tattctgctc 35400
 ctggttcacc tgcagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
 ctctccctc agcaataact gtaagtactc tttcatatcg tctccgaaat ttttagccat 35520
 aggaccccca ggaataagag aagggaagc cacattacag ataaaccgaa gtcccccca 35580
 gtgagcattg ccaaagttaa gattgaaata agcatgctgg ctagaccgg tgatatcttc 35640
 cagataactg gacagaaaat cgggtaagca atttttaaga aaatcaacaa aagaaaaatc 35700
 ttccagggtgc acgtttaggg cctcgggaac aacgatggag taagtgaag ggggtgcgttc 35760
 cagcatggtt agttagctga tctgtaaaaa aacaaaaaat aaaacattaa accatgctag 35820
 cctggcgaac aggtgggtaa atcgttctct ccagcaccag gcaggccacg ggggtctccg 35880
 cgcgaccctc gtaaaaattg tcgctatgat tgaaaaccat cacagagaga cgttcccggg 35940
 ggccggcgtg aatgattcga gaagaagcat acaccccg aacattggag tccgtgagtg 36000
 aaaaaaagcg gccgaggaag caatgaggca ctacaacgct cactctcaag tccagcaaag 36060
 cgatgccatg cggatgaagc acaaaatctt cagggtgcgt aaaaatgtaa ttactccct 36120
 cctgcacagg cagcgaagct cccgatccct ccagatacac atacaaagcc tcagcgtcca 36180
 tagcttaccg agcggcagca gcagcggcac acaacaggcg caagagtcag agaaaagact 36240
 gagctctaac ctgtccgccc gctctctgct caatatatag cccagatct aactgacgt 36300
 aaaggccaaa gtctaaaaat acccgccaaa taatcacaca cgcccagcac acgcccagaa 36360
 accggtgaca cactcagaaa aatacgcgca cttcctcaaa cggccaaact gccgtcattt 36420
 ccgggttccc acgctacgtc atcaaacac gactttcaaa ttccgtcgac cgttaaaaac 36480
 atacccgcc ccgcccctaa cggtcgccgc tccgcagcc aatcaccttc ctccctcccc 36540
 aaattcaaac agctcatttg catattaacg cgcacaaaaa gtttgaggta tattattgat 36600
 gatg 36604

<210> 3
 <211> 36535
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan7

<400> 3
 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaag agctgtttga 60

atttggggag ggaggaaggt gattggccga gagacgggcg accgttaggg gcggggcggg 120
 tgacgttttt aatacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
 ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggtgta 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
 gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
 gtggcgaccc tcctgagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgtatg 720
 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
 gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
 cctctctcca taccgccaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgca tcgaaccagg gagtgaagc tgcgggcgaa agcttttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
 acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactggttt 1200
 atttatgtat atgttttttt atgtgtaggc cccgtctctg acgtagatga gacccccact 1260
 tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
 agaccagttg cagtgaagat caccgggagg agagcagctg tggagagttt ggatgacttg 1380
 ctacaggggtg gggatgaacc tttggacttg tgtaccgga aacgccccag gcactaagtg 1440
 ccacacatgt gtgtttactt aagggtgatgt cagtatttat aggggtgtgga gtgcaataaa 1500
 atccgtgttg actttaagtg cgtggtttat gactcagggg tggggactgt gggatatataa 1560
 gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gacggtcttg 1620
 gaagactttc accagactag acagctgcta gagaactcat cggagggggg ctcttacctg 1680
 tggagattct gcttcgggtg gcctctagct aagctagtct atagggccaa acaggattat 1740
 aaggatcaat ttgaggatat tttgagagag tgtcctggta tttttgactc tctcaacttg 1800
 ggccatcagt ctacttttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc 1860
 agaactaccg ccgcggtagc cttttttgcc ttatccttg acaaatggag tcaagaaacc 1920
 catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg 1980
 tgccagcgcc tgaatgcaat ctccggctac ttgccagtac agccggtaga cacgctgagg 2040
 atcctgagtc tccagtcacc ccaggaacac caacgccgcc agcagccgca gcaggagcag 2100
 cagcaagagg aggaggagga tcgagaagag aaccgagag ccggtctgga ccctccggtg 2160

gcgaggaggagg aggagtagct gacttgtttc ccgagctgcg ccgggtgctg actaggtctt 2220
ccagtggacg ggagagggggg attaacggg agaggcatga ggagactagc cacagaactg 2280
aactgactgt cagtctgatg agccgcaggc gcccagaatc ggtgtggtgg catgaggttc 2340
agtcgcaggg gatagatgag gtctcggtga tgcatagaa atattccctg gaacaagtca 2400
agacttggtg gttggagcct gaggatgatt gggaggtagc catcaggaat tatgccaaagc 2460
tggctctgaa gccagacaag aagtacaaga ttaccaaact gattaatatc agaaattcct 2520
gctacatttc agggaaatggg gccgaggtgg agatcagtac ccaggagagg gtggccttca 2580
gatgttgtat gatgaatatg taccggggg tgggtggcat ggagggagtc acctttatga 2640
acgcgaggtt caggggtgat ggggtataatg ggggtggtctt tatggccaac accaagctga 2700
cagtgcacgg atgctccttc tttgggttca ataacatgtg catcgaggcc tggggcagtg 2760
tttcagttag gggatgcagc ttttcagcca actggatggg ggtcgtgggc agaaccaaga 2820
gcaaggtgtc agtgaagaaa tgcctgttcg agaggtgcca cctgggggtg atgagcgagg 2880
gcgaagccaa agtcaaacac tgcgcctcta ctgagacggg ctgctttgtg ctgatcaagg 2940
gcaatgccca agtcaagcat aacatgatct gtggggcctc ggatgagcgc ggctaccaga 3000
tgctgacctg cgccggtggg aacagccata tgctggccac cgtgcatgtg acctcgacc 3060
cccgaagac atggcccgag ttcgagcaca acgtcatgac ccgatgcaat gtgcacctgg 3120
ggtcccgcg aggcattgtc atgccctacc agtgcaacat gcaatttgtg aaggtgctgc 3180
tggagcccga tgccatgtcc agagttagcc tgacgggggt gtttgacatg aatgtggagc 3240
tgtggaaaat tctgagatat gatgaatcca agaccaggtg ccgggcctgc gaatgaggag 3300
gcaagcacgc caggcttcag cccgtgtgtg tggaggtgac ggaggacctg cgaccgatc 3360
atgtggtgtt gtcctgcaac gggacggagt tcggctccag cggggaagaa tctgactaga 3420
gtgagtagtg tttgggggag gtggagggct tgtatgaggg gcagaatgac taaaatctgt 3480
gtttttctgt gtgttcagc agcatgagc gaagcgctc ctttgaggga ggggtattca 3540
gcccttatct gacggggcgt ctccctcct gggcgggagt gcgtcagaat gtgatgggat 3600
ccacggtgga cgccgggccc gtgcagccc cgaactctc aaccctgacc tacgcgaccc 3660
tgagctcctc gtccgtggac gcagctgcc cgcagctgc tgcctccgc gccagcgccg 3720
tgcgcggaat ggccctggg gccggctact acagctctct ggtggccaac tcgacttcca 3780
ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg 3840
ccctgaccca gcgcctggg gagctgacct agcaggtggc tcagctgcag gcggagacgc 3900
gggccgcggt tgccacggtg aaaaccaaata aaaaatgaa tcaataaata aacgggagacg 3960
gttgttgatt ttaacacaga gtcttgaatc tttatttgat ttttcgcgcg cggtaggccc 4020
tggaaccacc gtctcgatca ttgagcacc ggtggatttt ttccaggacc cggtagaggt 4080
gggcttgat gttgaggtac atgggcatga gcccgctccg ggggtggagg tagctccatt 4140
gcagggcctc gtgctcggg gtggtgttgt aaatcaccca gtcatagcag gggcgagg 4200

cgtggtgctg cacgatgtcc ttgaggagga gactgatggc cacgggcagc cccttggtgt 4260
 aggtgttgac gaacctgttg agctgggagg gatgcatgcg gggggagatg agatgcatct 4320
 tggcctggat cttgagattg gcgatgttcc cgcccagatc ccgccggggg ttcatgttgt 4380
 gcaggaccac cagcacggtg tatccggtgc acttggggaa tttgtcatgc aacttggaag 4440
 ggaaggcgtg aaagaatttg gagacgccct tgtgaccgcc caggttttcc atgcactcat 4500
 ccatgatgat ggcgatggg ccgtgggcgg cggcctgggc aaagacgttt cgggggtcgg 4560
 acacatcgta gttgtggtcc tgggtgagct cgtcataggc cattttaatg aatttggggc 4620
 ggaggggtgc cgaactgggg acgaagggtgc cctcgatccc gggggcgtag ttgccctcgc 4680
 agatctgcat ctcccaggcc ttgagctcgg agggggggat catgtccacc tgcggggcga 4740
 tgaaaaaac ggtttccggg gcgggggaga tgagctgggc cgaaagcagg ttccggagca 4800
 gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc tgcaggtggt 4860
 agttgagga gagacagctg ccgtcctcgc ggaggagggg ggccacctcg ttcatcatct 4920
 cgcgcacatg catgttctcg cgacagagtt ccgccaggag gcgctcgccc cccagcgaga 4980
 ggagctcttg cagcgaggcg aagtttttca gcggcttgag yccgtcggcc atgggcattt 5040
 tggagagggg ctgttgcaag agttccagac ggtcccagag ctcggtgatg tgctctaggg 5100
 catctcgatc cagcagacct cctcgtttcg cgggttgggg cgactgcggg agtagggcac 5160
 caggcgatgg gcgtccagcg aggccagggt ccggtccttc cagggtcgca gggtcgcgt 5220
 cagcgtggtc tccgtcacgg tgaaggggtg cgcgccgggc tgggcgcttg cgagggtgcg 5280
 cttcaggctc atccggctgg tcgagaaccg ctcccggctg gcgccctgcg cgtcggccag 5340
 gtagcaattg agcatgagtt cgtagttgag gcctcggcc gcgtggccct tggcgcgag 5400
 cttacctttg gaagtgtgtc cgacagcggg acagaggagg gacttgaggg cgtagagctt 5460
 gggggcgagg aagacggact cgggggcgta ggcgtccgcg ccgcagctgg cgacagcgg 5520
 ctgcactcc acgagccagg tgaggctcggg ccggttgggg tcaaaaacga ggtttcctcc 5580
 gtgctttttg atgcgtttct tacctctggt ctccatgagc tcgtgtcccc gctgggtgac 5640
 aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cggctcctga gcggggtgcc 5700
 gcggtcctcg tcgtagagga accccgcca ctccgagacg aaggcccggg tccaggccag 5760
 cacgaaggag gccacgtggg aggggtagcg gtcgttgtcc accagcgggt ccacctctc 5820
 cagggtatgc aagcacatgt cccctcgtc cacatccagg aagggtgatt gcttgtaagt 5880
 gtagggcacg tgaccggggg tcccggccgg gggggtataa aagggggcg gcccctgctc 5940
 gtcctcactg tcttccggat cgctgtccag gagcgccagc tgttggggta ggtattccct 6000
 ctgaaggct ggcataacct cggcactcag gttgtcagtt tctagaaacg aggaggattt 6060
 gatattgacg gtgccgttg agacgccttt catgagcccc tcgtccatct ggtcagaaaa 6120
 gacgatcttt ttgttgcga gcttgggtggc gaaggagccg tagagggcgt tggagaggag 6180
 cttggcgatg gagcgcatgg tctggttctt ttccttgctg gcgcgctcct tggcggcgat 6240
 gttgagctgc acgtactcgc gcgccacga cttccattcg gggaagacgg tggtagctc 6300

gtcgggcacg attctgaccc gccagccgcg gttgtgcagg gtgatgaggt ccacgctggt	6360
ggccacctcg ccgcgcaggg gctcgttggt ccagcagagg cggccgccct tgcgcgagca	6420
gaaggggggc agcgggtcca gcatgagctc gtcggggggg tcggcgcca cgggtgaagat	6480
gccgggcaga agctcggggg cgaagtagct gatgcaggtg tccagatcgt ccagcgccgc	6540
ttgccagtcg cgcacggcca gcgcgcgctc gtaggggctg aggggcgtgc cccagggcat	6600
ggggtgcgtg agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga ggggctcctc	6660
gaggacgccg atgtaggtgg ggtagcagcg ccccccgcgg atgctggcg gcacgtagtc	6720
gtacagctcg tgcgagggcg cgaggagccc cgtgccgagg ttggagcgtt gcggcttttc	6780
ggcgcggtag acgatctggc ggaagatggc gtgggagttg gaggagatgg tgggcctctg	6840
gaagatgttg aagtgggctg ggggcaggcc gaccgagtcc ctgatgaagt gggcgtagga	6900
gtcctgcagc ttggcgacga gctcggcggg gacgaggacg tccagggcg agtagtcgag	6960
ggtctcttg atgatgtcgt acttgagctg gcccttctgc ttccacagct cgcggttgag	7020
aaggaactct tcgcggtcct tccagtactc ttcgaggggg aaccgcctc gatcggcacg	7080
gtaagagccc accatgtaga actggttgac ggccttgtag gcgcagcagc ctttctccac	7140
ggggagggcg taagcttggt cggccttgcg caggaggtg tgggtgaggg cgaaggtgtc	7200
gcgcaccatg accttgagga actggtgctt gaagtcgagg tcgtcgcagc cgccctgctc	7260
ccagagctgg aagtccgtgc gcttcttgta ggcgggggtg ggcaaagcga aagtaacatc	7320
gttgaagagg atcttgcccg cgcggggcat gaagttgcga gtgatgcgga aaggctgggg	7380
cacctcggcc cggttggtga tgacctgggc ggcgaggacg atctcgtcga agccgttgat	7440
gttgtgcccg acgatgtaga gttccacgaa tcgcgggcgg cccttaacgt ggggcagctt	7500
cttgagctcg tcgtaggtga gctcggcggg gtcgctgagc ccgtgctgct cgagggccca	7560
gtcggcgacg tgggggttg cgctgaggaa ggaagtccag agatccacgg ccagggcggg	7620
ctgcaagcgg tcccgtact gacggaactg ctggcccacg gccatTTTTT cgggggtgac	7680
gcagtagaag gtgcgggggt cgccgtgcc gcggtcccac ttgagctgga gggcgaggtc	7740
gtgggcgagc tcgacgagcg gcgggtcccc ggagagtttc atgaccagca tgaaggggac	7800
gagctgcttg ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg tgaggaagag	7860
cctttcgggt cgaggatgcg agccgatggg gaagaactgg atctcctgcc accagttgga	7920
ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgagc actcgtgctt	7980
gtgtttatac aagcgtccgc agtgctcgca acgctgcacg ggatgcacgt gctgcacgag	8040
ctgtacctgg gttcctttga cgaggaattt cagtgggcag tggagcgtg gcggctgcat	8100
ctggtgctgt actacgtcct ggccatcggc gtggccatcg tctgcctcga tgggtggtcat	8160
gctgacgagc ccgcgcggga ggcaggtcca gacttcggct cggacgggtc ggagagcgag	8220
gacgagggcg cgcagggcgg agctgtccag ggtcctgaga cgctgcggag tcaggtcagt	8280
gggcagcggc ggcgcgcggg tgacttgacg gagcttttcc agggcgcgcg ggaggtccag	8340

atggtacttg	atctccacgg	cgccgttggg	ggcgacgtcc	acggcttgca	gggtcccgtg	8400
cccctggggc	gccaccaccg	tgccccgttt	cttcttgggc	gctgcttcca	tgccggtcag	8460
aagcggcggc	gaggacgcgc	gccgggcggc	aggggcggct	cgggacccgg	aggcaggggc	8520
ggcaggggca	cgtcggcgcc	gcgcgcgggc	aggttctggg	actgcgcccg	gagaagactg	8580
gcgtgagcga	cgacgcgacg	gttgacgtcc	tggatctgac	gcctctgggt	gaaggccacg	8640
ggacccgtga	gtttgaacct	gaaagagagt	tcgacagaat	caatctcggt	atcgttgacg	8700
gcggcctgcc	gcaggatctc	ttgcacgtcg	cccagattgt	cctggtaggc	gatctcggtc	8760
atgaactgct	cgatctcctc	ctcctgaagg	tctccgcggc	cggcgcgctc	gacggtggcc	8820
gcgaggtcgt	tggagatgcg	gcccattgagc	tcgagaagg	cgttcatgcc	ggcctcgttc	8880
cagacgcggc	tgtagaccac	ggctccgtcg	gggtcgcgcg	cgcgcattgac	cacctgggcg	8940
aggttgagct	cgacgtggcg	cgtgaagacc	gcgtagttgc	agaggcgctg	gtagaggtag	9000
ttgagcgtgg	tggcgatgtg	ctcggtgacg	aagaagtaca	tgatccagcg	gcggagcggc	9060
atctcgtga	cgtcgcgccg	ggcttccaag	cgctccatgg	cctcgtagaa	gtccacggcg	9120
aagttgaaaa	actgggagtt	gcgcgccgag	acggtcaact	cctcctccag	aagacggatg	9180
agctcagcga	tgggtggcgcg	cacctcgcgc	tcgaaggccc	cgggggggctc	ctcttcttcc	9240
atctcttctt	cctccactaa	catctcttct	acttctctct	caggaggcg	cggcggggga	9300
ggggccctgc	gtcgcggcg	gcgcacgggc	agacggtcga	tgaagcgctc	gatggtctcc	9360
ccgcgccggc	gacgcattgt	ctcggtgacg	gcgcgcccg	cctcgcgggg	ccgcagcgtg	9420
aagacgccgc	cgcgcattct	cagggtggccg	ccgggggggt	ctccgttggg	caggagagg	9480
gcgctgacga	tgcatttat	caattggccc	gtagggactc	cgcgcaagga	cctgagcgtc	9540
tcgagatcca	cgggatccga	aaaccgctga	acgaaggctt	cgagccagtc	gcagtcgcaa	9600
ggtaggctga	gcccggtttc	ttgttcttcg	gggatttcgg	gaggcgggcg	ggcgatgctg	9660
ctggtgatga	agttgaagta	ggcggtcctg	agacggcgga	tgggtggcgag	gagcaccagg	9720
tccttgggcc	cggcttgctg	gatgcgcaga	cggtcggcca	tgccccaggc	gtggtcctga	9780
cacctggcga	ggtccttgta	gtagtcctgc	atgagccgct	ccacgggcac	ctcctcctcg	9840
cccgcgcggc	cgtgcatgcg	cgtgagcccc	aaccgcgct	ggggctggac	gagcgccagg	9900
tcggcgacga	cgcgctcggc	gaggatggcc	tgctgtatct	gggtgagggt	ggtctggaag	9960
tcgtcgaagt	cgacgaagcg	gtggtaggct	ccggtgttga	tggatatagga	gcagttggcc	10020
atgacggacc	agttgacggt	ctggtggccg	ggtcgcacga	gctcgtggta	cttgaggcgc	10080
gagtaggcgc	gcgtgtcgaa	gatgtagtcg	ttgcagggtgc	gcacgaggta	ctggtatccg	10140
acgaggaagt	gcggcgggcg	ctggcggtag	agcggccatc	gctcgggtggc	gggggcgcgc	10200
ggcgcgaggt	cctcgagcat	gaggcggtgg	tagccgtaga	tgtacctgga	catccagggtg	10260
atgccggcgg	cgggtggtgga	ggcgcgcggg	aactcgcgga	cgcggttcca	gatgttgcg	10320
agcggcagga	agtagttcat	ggtggccgcg	gtctggcccc	tgaggcgcgc	gcagtcgtgg	10380
atgctctaga	catacgggca	aaaacgaaag	cggtcagcgg	ctcgactccg	tggcctggag	10440

gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag 10500
ccgcagctaa cgtggtactg gcactcccg tctgacccaa gcctgctaac gaaacctcca 10560
ggatacggag gcgggtcgtt ttttggcctt ggtcgttgg catgaaaaac tagtaagcgc 10620
ggaaagcgac cggccgcgat ggctcgtgc cgtagtctgg agaaagaatc gccagggttg 10680
cgttgcggtg tgccccggtt cgagcctcag cgctcggcgc cggccggatt ccgcggctaa 10740
cgtgggcgtg gctgccccgt cgtttccaag accccttagc cagccgactt ctccagttac 10800
ggagcgagcc cctctttttt ttgtgttttt gccagatgca tcccgtactg cggcagatgc 10860
gccccaccc tccacctcaa ccgcccctac cgccgcagca gcagcaacag ccggcgcttc 10920
tgcccccgcc ccagcagcag ccagccacta ccgcggcggc cgccgtgagc ggagccggcg 10980
ttcagtatga cctggccttg gaagagggcg aggggctggc gcggctgggg gcgtcgtcgc 11040
cggagcggca ccgcgcgtg cagatgaaaa gggacgctcg cgaggcctac gtgccaagc 11100
agaacctgtt cagagacagg agcggcgagg agcccgagga gatgcgcgcc tcccgtttcc 11160
acgcggggcg ggagctgcgg cgcggcctgg accgaaagcg ggtgctgagg gacgaggatt 11220
tcgaggcgga cgagctgacg gggatcagcc ccgcgcgcgc gcacgtggcc gcggccaacc 11280
tggtcacggc gtacgagcag accgtgaagg aggagagcaa cttccaaaaa tccttcaaca 11340
accacgtgcg cacgctgatc gcgcgcgagg aggtgaccct gggcctgatg cacctgtggg 11400
acctgctgga ggccatcgtg cagaacccca cgagcaagcc gctgacggcg cagctgtttc 11460
tggtggtgca gcacagtcgg gacaacgaga cgttcaggga ggcgtgctg aatatcaccg 11520
agcccagagg ccgctggctc ctggacctgg tgaacattct gcagagcatc gtggtgcagg 11580
agcgcgggct gccgctgtcc gagaagctgg cggctatcaa cttctcgggt ctgagcctgg 11640
gcaagtacta cgctaggaag atctacaaga ccccgtagct gcccatagac aaggagggtga 11700
agatcgacgg gttttacatg cgcagtaccc tgaaagtgct gaccctgagc gacgatctgg 11760
gggtgtaccg caacgacagg atgcaccgcg cggtagcgc cagccgccgg cgcgagctga 11820
gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgaggggg 11880
agagctactt tgacatgggc gcggacctgc gctggcagcc cagccgccgg gccttggaag 11940
ctgccggcgg tttcccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000
tggaagactg atggcgcgac cgtatttttg ctagatgcag caacagccac cgcctcctga 12060
ttccgcgatg cgggcggcgc tgcagagcca gccgtccggc attaactcct cggacgattg 12120
gacctcaggcc atgcaacgca tcatggcgct gacgaccgc aatcccgaag ctttagaca 12180
gcagcctcag gccaaccggc tctcgccat cctggaggcc gtggtgccct cgcgctcgaa 12240
ccccacgcac gagaagggtg tggccatcgt gaacgcgctg gtggagaaca aggccatccg 12300
cggcgacgag gccgggctgg tgtacaacgc gctgctggag cgcggtggcc gctacaacag 12360
caccaacgtg cagacgaacc tggaccgat ggtgaccgac gtgcgcgagg cgggtgtcga 12420
gcgcgagcgg ttccaccgcg agtcgaacct gggctccatg gtggcgctga acgccttcct 12480

gagcacgcag cccgccaacg tgccccgggg ccaggaggac tacaccaact tcatcagcgc 12540
 gctgcggtg atggtggccg aggtgcccc gagcgaggtg taccagtcgg ggccggacta 12600
 cttcttccag accagtcgcc agggcttgca gaccgtgaac ctgagccagg ctttcaagaa 12660
 cttgcagggg ctgtggggcg tgcaggcccc ggtcggggac cgcgcgacgg tgtcagacct 12720
 gctgacgccg aactcgcgcc tgctgctgct gctggtggcg cccttcacgg acagcggcag 12780
 cgtgagccgc gactcgtacc tgggctacct gcttaacctg taccgcgagg ccatcgggca 12840
 ggcgcacgtg gacgagcaga cctaccagga gatcaccac gtgagccgcg cgctgggcca 12900
 ggaggacccg ggcaacctgg aggccaccct gaacttcctg ctgaccaacc ggtcgcagaa 12960
 gatcccgccc cagtacgcgc tgagcaccga ggaggagcgc atcctgcgt acgtgcagca 13020
 gagcgtgggg ctgttcctga tgcaggaggg ggccacgccc agcgccgcgc tcgacatgac 13080
 cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat 13140
 ggactacttg catcgggcg cgcctatgaa ctcggactac ttaccaacg ccatcttgaa 13200
 cccgcactgg ctcccgccgc ccgggttcta cacgggagag tacgacatgc ccgaccccaa 13260
 cgacgggttc ctgtgggacg acgtggacag cagcgtgttc tcgccgcgcc ccgccaccac 13320
 cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg gtcgcgcggg 13380
 tgctgcccg gcggtgcctg aggccgccag ccccttcccg agcctgccct tttcgctgaa 13440
 cagcgtgcgc agcagcgagc tgggtcggct gacgcggccg cgctgtctgg gcgaggagga 13500
 gtacctgaac gactccttgt tgaggccga gcgcgagaag aacttccca ataacgggat 13560
 agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc acagggacga 13620
 gccccgagct agcagcagcg caggcaccg tagacgccag cgacacgaca ggcagcgggg 13680
 tctggtgtgg gacgatgagg attccgccga cgacagcagc gtgttgact tgggtgggag 13740
 tgggtggtgt aaccggttcg ctacttgcg ccccgatatc gggcgctga tgaagaatc 13800
 tgaaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttctctgt 13860
 tgtttgtagt agtatgatga ggcgcgtgta cccggagggt cctcctccct cgtacgagag 13920
 cgtgatgcag caggcgggtg cggcggcgat gcagccccg ctggaggcgc cttacgtgcc 13980
 cccgcggtac ctggcgcta cggaggggag gaacagcatt cgttactcg agctggcacc 14040
 cttgtacgat accaccggt tgtacctggt ggacaacaag tcggcggaac tcgcctcgct 14100
 gaactaccag aacgaccaca gcaacttcct gaccaccgtg gtgcagaaca acgatttcac 14160
 cccacggag gccagcacc agaccatcaa ctttgacgag cgctcgcggt ggggcggcca 14220
 gctgaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa 14280
 gttcaaggcg cgggtgatgg tctcgcgcaa gaccccaat ggggtcgcgg tggatgagaa 14340
 ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccgaggg 14400
 caacttctcg gtgacctga ccatcgatct gatgaacaac gccatcatcg acaactactt 14460
 ggcggtgggg cgtcagaacg ggggtgctga gagcgacatc ggcgtgaagt tcgacacgcg 14520
 caacttccgg ctgggctggg acccgtgac cgagctggtg atgccgggag tgtacaccaa 14580

cgaggccttc caccgagaca tcgtcctgct gcccggctgc ggcgtggact tcaccgagag 14640
ccgcctcagc aacctgctgg gcatccgcaa gcggcagccc ttccaggagg gcttcagat 14700
cctgtacgag gacctggagg ggggcaacat ccccgcgctc ttggatgtcg aagcctatga 14760
gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccg cctctaccga 14820
ggtgcggggc gataattttg ctagcgccgc ggcagtggcc gaggcggctg aaaccgaaag 14880
taagatagtc atccagccgg tggagaagga cagcaaggac aggagctaca acgtgctcgc 14940
ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgaccccga 15000
gaagggcgtg cgctcctgga cgctgctcac cacctcggac gtcacctgcg gcgtggagca 15060
agtctactgg tcgctgcccc acatgatgca agacccggtc accttcgct ccacgcgtca 15120
agttagcaac taccgggtgg tgggcgcca gctcctgccc gtctactcca agagcttctt 15180
caacgagcag gccgtctact cgcagcagct gcgcgccttc acctcgtca cgcacgtctt 15240
caaccgcttc cccgagaacc agatcctcgt ccgcccgcgc gcgcccacca ttaccaccgt 15300
cagtgaanaa gttcctgctc tcacagatca cgggaccctg ccgctgcgca gcagtatccg 15360
gggagtccag cgctgaccg tctactgacg cagacgccgc acctgcccct acgtctacaa 15420
ggccctgggc gtagtcgctc cgcgctcct ctcgagccgc accttctaaa aaatgtccat 15480
tctcatctcg cccagtaata acaccggtt gggcctgcgc gcgcccagca agatgtacgg 15540
aggcgtcgc caacgctcca cgcaacacc cgtgcgcgtg cgcgggcact tccgcgctcc 15600
ctggggcgcc ctcaagggcc gcgtgcgctc gcgcaccacc gtcgacgacg tgatcgacca 15660
ggtggtggcc gacgcgcgca actacacgcc cgccgccgcg cccgcctcca ccgtggacgc 15720
cgtcatcgac agcgtggtgg ccgatgcgcg ccggtacgcc cgcgccaaga gccggcggcg 15780
gcgcatcgcc cggcggcacc ggagcaccgc cgccatgcgc gcggcgcgag ccttgctgcg 15840
cagggccagg cgacggggac gcagggccat gctcagggcg gccagacgcg cggcctccgg 15900
cagcagcagc gccggcagga cccgcagacg cgcgccacg gcggcggcg cggccatcgc 15960
cagcatgtcc cgcccgcggc gcggcaacgt gtactgggtg cgcgacgccg ccaccggtgt 16020
gcgcgtgccc gtgcgcaccc gccccctcg cacttgaaga tgctgacttc gcgatgttga 16080
tgtgtcccag cggcgaggag gatgtccaag cgcaataca aggaagagat gctccaggtc 16140
atcgcgctg agatctacgg cccgcggtg aaggaggaaa gaaagccccg caaactgaag 16200
cgggtcaaaa aggacaaaaa ggaggaggaa gatgtggacg gactggtgga gtttgtgcgc 16260
gagttcgccc cccggcggcg cgtgcagtgg cgcgggcgga aagtgaacc ggtgctgcg 16320
ccggcacca cgggtgtctt cagcccggc gagcgttccg gctccgcctc caagcgtcc 16380
tacgacgagg tgtacgggga cgaggacatc ctcgagcagg cggtcgagcg tctgggagag 16440
tttgcttacg gcaagcgcag ccgccccgcg cccttgaaag aggaggcggg gtccatcccg 16500
ctggaccacg gcaacccac gccgagcctg aagccggtga ccctgcagca ggtgctgcg 16560
agcgcggcgc cgcgccggg cttcaagcgc gagggcggcg aggatctgta cccgaccatg 16620

cagctgatgg tgcccaagcg ccagaagctg gaggacgtgc tggagcacat gaaggtggac 16680
cccagagtg c agcccgaggt caaggtgcgg cccatcaagc aggtggcccc gggcctgggc 16740
gtgcagaccg tggacatcaa gatccccacg gagcccatgg aaacgcagac cgagcccgtg 16800
aagcccagca ccagcaccat ggaggtgcag acggatccct ggatgccggc gccggcttcc 16860
accactcgcc gaagacgcaa gtacggcgcg gccagcctgc tgatgcccaa ctacgcgctg 16920
catccttcca tcatccccac gccgggctac cgcggcacgc gcttctaccg cggctacacc 16980
agcagccgcc gcaagaccac caccgcgcg cgcgctcgtc gcacccgccg cagcagcacc 17040
gcgacttccg ccgccgcctt ggtgcggaga gtgtaccgca gcgggcgcga gcctctgacc 17100
ctgccgcg cgcgctacca cccgagcatc gccatttaac tctgccgtcg cctcctactt 17160
gcagatatgg ccctcacatg ccgcctccgc gtccccatta cgggctaccg aggaagaaaag 17220
ccgcgccgta gaaggctgac ggggaacggg ctgctgcgcc atcaccaccg gcggcgggcg 17280
gccatcagca agcgggttggg gggaggcttc ctgccgcgc tgatcccat catcgccgcg 17340
gcgatcgggg cgatccccg catagcttcc gtggcggtgc aggcctctca gcgccactga 17400
gacacagctt ggaaaatttg taataaaaa atggactgac gctcctggtc ctgtgatgtg 17460
tgtttttaga tggaagacat caatttttcg tccctggcac cgcgacacgg cacgcggccg 17520
tttatgggca cctggagcga catcggaac agccaactga acggggggcg cttcaattgg 17580
agcagtctct ggagcgggct taagaatttc ggggtccacgc tcaaaccta tggcaacaag 17640
gcgtggaaca gcagcacagg gcaggcgctg agggaaaagc tgaaagagca gaacttccag 17700
cagaaggtgg tcgatggcct ggcctcgggc atcaacgggg tgggtggacct ggccaaccag 17760
gccgtgcaga aacagatcaa cagccgcctg gacgcggtcc cggccgcggg gtccgtggag 17820
atgccccagg tggaggagga gctgcctccc ctggacaagc gcggcgacaa gcgaccgcgt 17880
cccagcgcgg aggagacgct gctgacgcac acggacgagc cgcccccgta cgaggaggcg 17940
gtgaaactgg gtctgcccac cacgcggccc gtggcgcttc tggccaccgg ggtgctgaaa 18000
cccagcagca gcagccagcc cgcgaccctg gacttgcttc cgcctgcttc ccgcccctcc 18060
acagtggcta agcccctgcc gccggtggcc gtcgctcgc gcgcccccg aggccgcccc 18120
caggcgaact ggcagagcac tctgaacagc atcgtgggtc tgggagtga gagtgtgaag 18180
cgccgccgct gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt 18240
atgtccgccg accagaagga ggaagaggcg cgtcgccgag ttgcaagatg gccaccccat 18300
cgatgctgcc ccagtgggcg tacatgcaca tcgccggaca ggacgcttcg gagtacctga 18360
gtccgggtct ggtgcagttc gccgcgccca cagacaccta cttcagtctg gggacaaga 18420
ttaggaacct cacggtggcg cccacgcagc atgtgaccac cgaccgcagc cagcggctga 18480
cgctgcgctt cgtgcccgtg gaccgcgagg acaacaccta ctcgtacaaa gtgcgctaca 18540
cgctggccgt gggcgacaac cgcgtgctgg acatggccag cacctacttt gacatccgcg 18600
gcgtgctgga tcggggggccc agcttcaaac cctactccgg caccgcctac aacagcctgg 18660
ctccaaggg agcgcccaac acttgccagt ggacatataa agctgggtgat actgatacag 18720

aaaaaaccta	tacatatgga	aatgcacctg	tgcaaggcat	tagcattaca	aaggatggta	18780
ttcaacttgg	aactgacagc	gatggtcagg	caatctatgc	agacgaaact	tatcaaccag	18840
agcctcaagt	gggtgatgct	gaatggcatg	acatcactgg	tactgatgaa	aaatatggag	18900
gcagagctct	taagcctgac	accaaaatga	agccttgcta	tggttctttt	gccaagccta	18960
ccaataaaga	aggaggccag	gcaaattgtga	aaaccgaaac	aggcgggtacc	aaagaatatg	19020
acattgacat	ggcatttcttc	gataatcgaa	gtgcagctgc	cgccggccta	gccccagaaa	19080
ttgttttgta	tactgagaat	gtggatctgg	aaactccaga	tacccatatt	gtatacaagg	19140
caggtacaga	tgacagtagc	tcttctatca	atttggggtca	gcagtccatg	cccaacagac	19200
ccaactacat	tggcttcaga	gacaacttta	tcggtctgat	gtactacaac	agcactggca	19260
atatgggtgt	actggctgga	caggcctccc	agctgaatgc	tgtggtggac	ttgcaggaca	19320
gaaacaccga	actgtcctac	cagctcttgc	ttgactctct	gggtgacaga	accaggtatt	19380
tcagtatgtg	gaatcaggcg	gtggacagtt	atgaccccga	tgtgcgcat	attgaaaatc	19440
acggtgtgga	ggatgaactt	cctaactatt	gcttccccct	ggatgctgtg	ggtagaactg	19500
atacttacca	gggaattaag	gccaatggtg	ataatcaaac	cacctggacc	aaagatgata	19560
ctgttaatga	tgctaataa	ttgggcaagg	gcaatccttt	cgccatggag	atcaacatcc	19620
aggccaacct	gtggcggaac	ttcctctacg	cgaacgtggc	gctgtacctg	cccgactcct	19680
acaagtacac	gccggccaac	atcacgtgc	ccaccaacac	caacacctac	gattacatga	19740
acggccgcgt	ggtggcgccc	tcgctggtgg	acgcctacat	caacatcggg	gcgcgctggg	19800
cgctggaccc	catggacaac	gtcaaccctt	tcaaccacca	ccgcaacgcg	ggcctgcgat	19860
accgctccat	gctcctgggc	aacgggcgct	acgtgccctt	ccacatccag	gtgccccaaa	19920
agtttttcgc	catcaagagc	ctcctgctcc	tgcccggttc	ctacacctac	gagtggaaact	19980
tccgcaagga	cgtcaacatg	atcctgcaga	gctccctcgg	caacgacctg	cgcacggacg	20040
gggcctccat	cgccttcacc	agcatcaacc	tctacgccac	cttcttcccc	atggcgcaca	20100
acaccgcctc	cacgctcgag	gccatgctgc	gcaacgacac	caacgaccag	tccttcaacg	20160
actacctctc	ggcggccaac	atgctctacc	ccatcccggc	caacgccacc	aacgtgccca	20220
tctccatccc	ctcgcgcaac	tgggccgcct	tccgcggctg	gtccttcacg	cgcctcaaga	20280
cccgcgagac	gccctcgtc	ggctccgggt	tcgacccta	cttcgtctac	tcgggctcca	20340
tcccctacct	cgacggcacc	ttctacctca	accacacctt	caagaagggtc	tccatcacct	20400
tcgactcctc	cgtcagctgg	cccggcaacg	accgcctcct	gacgccaac	gagttcgaaa	20460
tcaagcgac	cgtcgacgga	gaggggtaca	acgtggccca	gtgcaacatg	accaaggact	20520
ggttcctggg	ccagatgctg	gcccactaca	acatcggcta	ccagggttc	tacgtgcccc	20580
agggctacaa	ggaccgcatg	tactccttct	tccgcaactt	ccagcccatg	agccgccagg	20640
tcgtggacga	ggtcaactac	aaggactacc	aggccgtcac	cctggcctac	cagcacaaca	20700
actcgggctt	cgtcggctac	ctcgcgcca	ccatgcgcca	gggccagccc	taccccgcca	20760

actacccta cccgctcatc ggcaagagcg ccgtcgccag cgtcaccag aaaaagttcc 20820
tctgcgaccg ggtcatgtgg cgcacccct tctccagcaa cttcatgtcc atgggcgcg 20880
tcaccgacct cggccagaac atgctctacg ccaactccgc ccacgcgcta gacatgaatt 20940
tcgaagtcga ccccatggat gagtccaccc ttctctatgt tgtcttcgaa gtcttcgacg 21000
tcgtccgagt gcaccagccc caccgcggcg tcatcgaggc cgtctacctg cgcacgccct 21060
tctcggccgg caacgccacc acctaagcct cttgcttctt gcaagatgac ggctgcgcg 21120
ggctccggcg agcaggagct cagggccatc ctccgcgacc tgggctgcg ggctgcttc 21180
ctgggcacct tcgacaagcg cttcccggga ttcatggccc cgcacaagct ggctgcgcc 21240
atcgtcaaca cggccggccg cgagaccggg ggcgagcact ggctggcctt cgcctggaac 21300
ccgcgctccc acacctgcta cctcttcgac cccttcgggt tctcggacga gcgcctcaag 21360
cagatctacc agttcgagta cgagggcctg ctgctcgca gcgccctggc caccgaggac 21420
cgctgctca ccctggaaaa gtccaccag accgtgcagg gtccgcgctc ggccgcctgc 21480
gggctcttct gctgcatgtt cctgcacgcc ttctgctact ggcccgaccg ccccatggac 21540
aagaaccca ccatgaactt gctgacgggg gtgcccacg gcatgctcca gtcgccccag 21600
gtggaacca ccctgcgccg caaccaggag gcgctctacc gcttctcaa cgccactcc 21660
gcctactttc gctcccaccg cgcgcgcatc gagaaggcca ccgccttcga ccgcatgaat 21720
caagacatgt aatccggtgt gtgtatgtga atgctttatt catcataata aacagcacat 21780
gtttatgcca cttctcttga ggctctgact ttatttagaa atcgaagggg ttctgccggc 21840
tctcggcatg gcccgcgggc agggatacgt tgcggaactg gtacttgggc agccacttga 21900
actcggggat cagcagcttc ggcacgggga ggtcggggaa cgagtcgctc cacagcttgc 21960
gcgtgagttg cagggcgccc agcaggtcgg gcgcggagat cttgaaatcg cagttgggac 22020
ccgcgttctg cgcgcgagag ttacggtaca cggggttgca gactggaac accatcaggg 22080
ccgggtgctt cacgctcgcc agcaccgtcg cgtcgggtgat gccctccacg tccagatcct 22140
cggcgttggc catcccgaag ggggtcatct tgcaggtctg ccgccccatg ctgggcacgc 22200
agccgggctt gtggttgcaa tcgcagtga gggggatcag catcatctgg gcctgctcgg 22260
agctcatgcc cgggtacatg gccttcatga aagcctccag ctggcggaag gcctgctcgg 22320
ccttgccgcc ctcggtgaag aagacccgc aggacttgct agagaactgg ttggtggcg 22380
agccagcgtc gtgcacgcag cagcgcgct cgttggtggc cagctgcacc acgctgcgcc 22440
cccagcggtt ctgggtgatc ttggcccggg cggggttctc cttcagcgcg cgctgcccgt 22500
tctcgctcgc cacatccatc tcgatcgtgt gctccttctg gatcatcacg gtcccgtgca 22560
ggcaccgcag cttgccctcg gcctcggtgc acccgtgcag ccacagcgcg cagccgggtgc 22620
tctcccagtt cttgtggcg atctgggagt gcgagtgac gaagccctgc aggaagcggc 22680
ccatcatcgt ggtcagggtc ttgttgctgg tgaaggtcag cggaatgccg cgggtgctcct 22740
cgttcacata caggtggcag atacggcggt acacctcgcc ctgctcgggc atcagctgga 22800
aggcggactt caggtcgctc tccacgcggg accggtccat cagcagcgtc atcacttcca 22860

tgcccttctc ccaggccgaa acgatcggca ggctcagggg gttcttcacc gttgtcatct 22920
tagtcgccgc cgccgaagtc aggggggtcgt tctcgtccag ggtctcaaac actcgtttgc 22980
cgtccttctc ggtgatgcgc acggggggaa agctgaagcc cacggccgcc agctcctcct 23040
cggcctgcct ttcgtcctcg ctgtcctggc tgatgtcttg caaaggcaca tgcttgggtct 23100
tgcggggttt ctttttgggc ggcagaggcg gcggcgagaga cgtgctgggc gagcgcgagt 23160
tctcgtcac cacgactatt tcttctcctt ggccgtcgtc cgagaccacg cggcggtagg 23220
catgcctctt ctggggcaga ggcggaggcg acgggctctc gcggttcggc gggcggttg 23280
cagagccccct tccgcgttcg ggggtgcgct cctggcggcg ctgctctgac tgacttcctc 23340
cgcgccggc cattgtgttc tcctagggag caagcatgga gactcagcca tcgtcgccaa 23400
catgccatc tgccccgcc gccgccgacg agaaccagca gcagcagaat gaaagcttaa 23460
ccgccccgcc gccagcccc acctccgacg ccgcagcccc agacatgcaa gagatggagg 23520
aatccatcga gattgacctg ggctacgtga cgcgcggga gcacgaggag gagctggcag 23580
cgcgcttttc agccccgaa gagaaccacc aagagcagcc agagcaggaa gcagagagcg 23640
agcagaacca ggctgggctc gagcatggcg actacctgag cggggcagag gacgtgctca 23700
tcaagcatct ggcccgccaa tgcacatcg tcaaggacgc gctgctcgac cgcgccgagg 23760
tgcccctcag cgtggcgag ctacgccgcg cctacgagcg caacctcttc tcgccgcgcg 23820
tgcccccaa gcgccagccc aacggcacct gcgagccaa cccgcgctc aacttctacc 23880
cggctcttcgc ggtgcccag gccctggcca cctaccacct ctttttcaag aaccaaagga 23940
tccccgtctc ctgccgcgcc aaccgcaccc gcgccgacgc cctgctcaac ctgggccccg 24000
gcgccgcct acctgatata gcctccttg aagaggttcc caagatcttc gagggctctg 24060
gcagcgacga gactcgggccc gcgaacgctc tgcaagggaag cggagaggag catgagcacc 24120
acagcgccct ggtggagttg gaaggcgaca acgcgcgcct ggcggtcctc aagcgcacgg 24180
tcgagctgac ccacttcgcc taccggcgcc tcaacctgcc cccaaggctc atgagcgccg 24240
tcatggacca ggtgctcatc aagcgcgctc cgcctctc ggaggaggag atgcaggacc 24300
ccgagagctc ggacgagggc aagcccgtgg tcagcgacga gcagctggcg cgctgggttg 24360
gagcgagtag cccccccag agcctggaag agcggcgcaa gctcatgatg gccgtggtcc 24420
tggtgaccgt ggagctggag tgtctgcgcc gtttcttcgc cgacgcggag accctgcgca 24480
aggtcgagga gaacctgcac tacctcttca gacacgggtt cgtgcgccag gcctgcaaga 24540
tctccaacgt ggagctgacc aacctggtct cctacatggg catcctgcac gagaaccgcc 24600
tggggcagaa cgtgctgcac accaccctgc gcggggaggc ccgccgcgac tacatccgcg 24660
actgctcta cctgtacctc tgccacacct ggcagacggg catgggcgtg tggcagcagt 24720
gcctggagga gcagaacctg aaagagctct gcaagctcct gcagaagaac ctcaaggccc 24780
tgtggaccgg gttcgacgag cgcaccaccg ccgcggacct ggccgacctc atcttccccg 24840
agcgctgcg gctgacgctg cgcaacgggc tgcccgaact tatgagccaa agcatgttgc 24900

aaaacttttcg ctcttttcac ctcgaacgct ccgggatcct gcccgccacc tgctccgcgc 24960
 tgccctcgga cttcgtgccg ctgaccttcc gcgagtggcc cccgccgctc tggagccact 25020
 gctacctgct gcgcctggcc aactacctgg cctaccactc ggacgtgatc gaggacgtca 25080
 gcggcgaggg cctgctcgag tgccactgcc gctgcaacct ctgcacgccg caccgctccc 25140
 tggcctgcaa cccccagctg ctgagcgaga ccagatcat cggcaccttc gagttgcaag 25200
 gccccggcga gggcaagggg ggtctgaaac tcaccccggg gctgtggacc tcggcctact 25260
 tgcgcaagtt cgtgcccag gactaccatc ctttcgagat caggttctac gaggaccaat 25320
 cccagccgcc caaggccgag ctgtcggcct gcgtcatcac ccagggggcc atcctggccc 25380
 aattgcaagc catccagaaa tcccgccaag aatttctgct gaaaaagggc cacggggtct 25440
 acttggaacc ccagaccgga gaggagctca accccagctt ccccaggat gccccgagga 25500
 agcagcaaga agctgaaagt ggagctgccg ccgccgccgg aggatttgga ggaagactgg 25560
 gagagcagtc aggcagagga ggaggagatg gaagactggg acagcactca ggcagaggag 25620
 gacagcctgc aagacagtct ggaggaggaa gacgaggtgg aggaggcaga ggaagaagca 25680
 gccgccgcca gaccgtcgtc ctccggcgag gaggagaaag caagcagcac ggataccatc 25740
 tccgctccgg gtcgggggtc cgccggccgg gcccacagta gatgggacga gaccgggcgc 25800
 ttcccgaacc ccaccacca gaccggtaag aaggagcggc agggatacaa gtcctggcgg 25860
 gggcacaaaa acgccatcgt ctctgcttg caagcctgcg ggggcaacat ctcttcacc 25920
 cggcgctacc tgctcttcca ccgcgggggtg aacttcccc gcaacatctt gcattactac 25980
 cgtcacctcc acagccccta ctactgtttc caagaagagg cagaaacca gcagcagcag 26040
 cagcagcaga aaaccagcgg cagcagctag aaaatccaca gcggcggcag gtggactgag 26100
 gatcgcgcg aacgagccgg cgcagaccgg ggagctgagg aaccggatct tccccaccct 26160
 ctatgccatc ttccagcaga gtcgggggca agagcaggaa ctgaaagtca agaaccgttc 26220
 tctgcgtcg ctacccgca gttgtctgta tcacaagagc gaagaccaac ttcagcgcac 26280
 tctcgaggac gccgaggctc tcttcaacaa gtactgcgcg ctactctta aagagtagcc 26340
 cgcgcccgcc cacacacgga aaaaggcggg aattacgtca ccacctgcgc ctttcgcccg 26400
 accatcatca tgagcaaaga gattcccacg cttacatgt ggagctacca gcccagatg 26460
 ggcttgccg ccggcgccgc ccaggactac tccaccgca tgaactggct cagtgccggg 26520
 cccgcgatga tctcacgggt gaatgacatc cgcgccacc gaaaccagat actcctagaa 26580
 cagtcagcga tcaccgccac gcccgcctat caccttaatc cgcgtaattg gcccgcgcc 26640
 ctggtgtacc aggaattcc ccagcccacg accgtactac ttccgcgaga cggccaggcc 26700
 gaagtccagc tgactaactc aggtgtccag ctggccggcg gcgccgcct gtgtcgtcac 26760
 cggccgctc agggataaaa gcggctggtg atccgaggca gaggcacaca gctcaacgac 26820
 gaggtggtga gctcttcgct gggctctgca cctgacggag tcttccaact cgccggatcg 26880
 gggagatctt cttcacgcc tcgtcaggcc gtcctgactt tggagagttc gtcctcgag 26940
 ccccgctcgg gtggcatcgg cactctccag ttcgtggagg agttcactcc ctccgtctac 27000

ttcaaccctt	tctccggctc	ccccggccac	taccgggacg	agttcatccc	gaacttcgac	27060
gccatcagcg	agtcggtgga	cggctacgat	tgaatgtccc	atggtggcgc	ggctgacctt	27120
gctcggcttc	gacacctgga	ccactgcccgc	cgcttccgct	gcttcgctcg	ggatctcgcc	27180
gagtttgctt	actttgagct	gcccaggag	caccctcagg	gcccggccca	cggagtgcgg	27240
atcgtcgtcg	aagggggtct	cgactcccac	ctgcttcgga	tcttcagcca	gcgtccgatc	27300
ctggccgagc	gcgagcaagg	acagaccctt	ctgaccctgt	actgcatctg	caaccacccc	27360
ggcctgcatg	aaagtctttg	ttgtctgctg	tgtactgagt	ataataaaag	ctgagatcag	27420
cgactactcc	ggacttccgt	gtgttcctgc	tatcaaccag	tccctgttct	tcaccgggaa	27480
cgagaccgag	ctccagctcc	agtgtgaagc	ccacaagaag	tacctcacct	ggctgttcca	27540
gggctctccg	atcgccgttg	tcaaccactg	cgacaacgac	ggagtcctgc	tgagcggccc	27600
tgccaacctt	actttttcca	cccgagaaag	caagctccag	ctcttccaac	ccttcctccc	27660
cgggacctat	cagtgcgtct	cgggaccctg	ccatcacacc	ttccacctga	tcccgaatac	27720
cacagcgtcg	ctccccgcta	ctaacaacca	aactaccac	caacgccacc	gtcgcgacct	27780
ttcctctggg	tctaatacca	ctaccggagg	tgagctccga	ggtcgaccaa	cctctgggat	27840
ttactacggc	ccctgggagg	tggtaggggt	aatagcgcta	ggcctagtgt	cgggtgggct	27900
tttggtcttc	tgctacctat	acctcccttg	ctgttcgtac	ttagtgggtg	tgtgttgctg	27960
gtttaagaaa	tggggaagat	caccctagt	agctgcggtg	tgctgggtgg	ggtgggtgct	28020
tcgattgtgg	gactgggagg	cgcggctgta	gtgaaggaga	aggccgatcc	ctgcttgcat	28080
ttcaatcccc	acaaatgcca	gctgagtttt	cagcccgatg	gcaatcgggtg	cgcgggtgctg	28140
atcaagtgcg	gatgggaatg	cgagaacgtg	agaatcgagt	acaataacaa	gactcggaac	28200
aatactctcg	cgtccgtgtg	gcagcccggg	gaccccgagt	ggtacaccgt	ctctgtcccc	28260
gggtgctgacg	gctccccgcg	caccgtgaat	aatactttca	tttttgcgca	catgtgcgac	28320
acggctcatgt	ggatgagcaa	gcagtacgat	atgtggcccc	ccacgaagga	gaacatcgtg	28380
gtctttctcca	tcgcttacag	cgtgtgcacg	gcgctaatac	ccgctatcgt	gtgcctgagc	28440
attcacatgc	tcacgctat	tcgccccaga	aataatgccg	aaaaagaaaa	acagccataa	28500
cacgtttttt	cacacacctt	tttcagacca	tggcctctgt	taaatttttg	cttttatttg	28560
ccagtctcat	tgccgtcatt	catggaatga	gtaatgagaa	aattactatt	tacactggca	28620
ctaatacacac	attgaaaggt	ccagaaaaag	ccacagaagt	ttcatgggtat	tgttatttta	28680
atgaatcaga	tgtatctact	gaactctgtg	gaaacaataa	caaaaaaat	gagagcatta	28740
ctctcatcaa	gtttcaatgt	ggatctgact	taaccctaata	taacatcact	agagactatg	28800
taggtatgta	ttatggaact	acagcaggca	tttcggacat	ggaattttat	caagtttctg	28860
tgtctgaacc	caccacgcct	agaatgacca	caaccacaaa	aactacacct	gttaccacta	28920
tacagctcac	taccaatggc	tttcttgcca	tgcttcaagt	ggctgaaaat	agcaccagca	28980
ttcaaccac	cccaccaggt	gaggaaattc	ccagatccat	gattggcatt	attgttgctg	29040

tagtggtgtg catgttgatc atcgcccttgt gcatggtgta ctatgccttc tgctacagaa 29100
 agcacagact gaacgacaag ctggaacact tactaagtgt tgaattttta ttttttagaa 29160
 ccatgaagat cctaggcctt ttagtttttt ctatcattac ctctgctcta tgcaattctg 29220
 acaatgagga cgttactgtc gttgtcggat caaattatac actaaaaggt ccagcaaaag 29280
 gtatgctttc gtggtattgt tgggtcggaa ctgacgagca acagacagaa ctttgcaatg 29340
 ctcaaaaagg caaaacctca aattctaaaa tctctaatta tcaatgcaat ggactgact 29400
 tagtattgct caatgtcacg aaagcatatg ctggcagtta cacctgccct ggagatgatg 29460
 ccgacaatat gatttttttac aaagtggag tggttgatcc cactactcca ccgcccacca 29520
 ccacaactac tcataccaca cacacagaac aaacaccaga ggcagcagaa gcagagttgg 29580
 ccttcagggt tcacggagat tcctttgctg tcaatacccc tacacccgat cagcgggtgtc 29640
 cggggctgct cgtcagcggc attgtcgggtg tgctttcggg attagcagtc ataatcatct 29700
 gcatgttcat ttttgcttgc tgctatagaa ggctttaccg acaaaaatca gacccactgc 29760
 tgaacctcta tgtttaattt tttccagagc catgaaggca gttagcgtc tagttttttg 29820
 ttctttgatt ggcattgttt ttagtgctgg gtttttgaaa aatcttacca tttatgaagg 29880
 tgagaatgcc actctagtgg gcatcagtgg tcaaatgtc agctggctaa aataccatct 29940
 agatgggtgg aaagacattt gcgattggaa tgtcactgtg tatacatgta atggagttaa 30000
 cctcaccatt actaatgcca cccaagatca gaatggtagg ttttaaggcc agagtttcac 30060
 tagaaataat ggggtatgaat ccataacat gtttatctat gacgtcactg tcatcagaaa 30120
 tgagactgcc accaccacac agatgcccac tacacacagt tctaccacta ctaccatgca 30180
 aaccacacag acaaccacta catcaactca gcatatgacc accactacag cagcaaagcc 30240
 aagtagtgca ggcctcagc cccaggcttt ggctttgaaa gctgcacaac ctagtacaac 30300
 tactaggacc aatgagcaga ctactgaatt tttgtccact gtcgagagcc acaccacagc 30360
 tacctccagt gccttctcta gcaccgcaa tctctcctcg ctttctctta caccaatcag 30420
 tcccgtact actcccaccc cagctcttct cccactccc ctgaagcaaa ctgaggacag 30480
 cggcatgcaa tggcagatca ccctgctcat tgtgatcggg ttggtcatcc tggccgtggt 30540
 gctctactac atcttctgcc gccgcattcc caacgcgcac cgcaaaccgg cctacaagcc 30600
 catcgttatc gggcagccgg agccgcttca ggtggaaggg ggtctaagga atcttctctt 30660
 ctcttttaca gtatggtgat tgaactatga ttcctagaca attcttgatc actattctta 30720
 tctgcctcct ccaagtctgt gccaccctcg ctctggtggc caacgccagt ccagactgta 30780
 ttggggccctt cgctcctac gtgctctttg ctttcatcac ctgcatctgc tgctgtagca 30840
 tagtctgcct gcttatcacc ttcttccagt tcattgactg gatctttgtg cgcacgcct 30900
 acctgcgcca ccaccccag taccgcgacc agcgagtggc gcggctgctc aggctcctct 30960
 gataagcatg cgggctctgc tacttctcgc gcttctgctg ttagtgctcc cccgccccgt 31020
 cgacccccgg tccccactc agtccccga agaggtccgc aaatgcaat tccaagaacc 31080
 ctggaaattc ctcaaatgct accgcaaaa atcagacatg cttcccagct ggatcatgat 31140

cattgggatc	gtgaacattc	tggcctgcac	cctcatctcc	tttgtgattt	acccctgctt	31200
tgacttttgg	tggaactcgc	cagaggcgct	ctatctcccg	cctgaacctg	acacaccacc	31260
acagcaacct	caggcacacg	cactaccacc	accacagcct	aggccacaat	acatgccccat	31320
attagactat	gaggccgagc	cacagcgacc	catgctcccc	gctattagtt	acttcaatct	31380
aaccggcgga	gatgactgac	ccactggcca	acaacaacgt	caacgacctt	ctcctggaca	31440
tggacggccg	cgcctcggag	cagcgactcg	cccaacttcg	cattcgccag	cagcaggaga	31500
gagccgtcaa	ggagctgcag	gacggcatag	ccatccacca	gtgcaagaaa	ggcatcttct	31560
gcctggtgaa	acaggccaag	atctcctacg	aggtcacccc	gaccgaccat	cgcctctcct	31620
acgagctcct	gcagcagcgc	cagaagttca	cctgcctggg	cggagtcaac	cccatcgtca	31680
tcaccagca	gtcgggcgat	accaaggggt	gcatccactg	ctcctgcgac	tcccccgact	31740
gcgtccacac	tctgatcaag	accctctgcg	gcctccgcga	cctcctcccc	atgaactaat	31800
caccccttta	tccagtgaag	taaatatcat	attgatgatg	atttaaataa	aaaataatca	31860
tttgatttga	aataaagata	caatcatatt	gatgatttga	gttttaaaaa	ataaagaatc	31920
acttacttga	aatctgatac	caggtctctg	tccatgtttt	ctgccaacac	cacctcactc	31980
ccctcttccc	agctctggta	ctgcagaccc	cggcgggctg	caaacttcct	ccacacgctg	32040
aaggggatgt	caaattcctc	ctgtccctca	atcttcattt	tatcttctat	cagatgtcca	32100
aaaagcgcgt	ccgggtggat	gatgacttcg	accccgctca	cccctacgat	gcagacaacg	32160
caccgaccgt	gcccttcac	aacccccct	tcgtctcttc	agatggattc	caagagaagc	32220
ccctgggggt	gctgtccctg	cgactggctg	accccgtcac	caccaagaac	ggggaaatca	32280
ccctcaagct	gggagagggg	gtggacctcg	actcctcggg	aaaactcatc	tccaacacgg	32340
ccaccaaggc	cgccgcccct	ctcagttttt	ccaacaacac	catttccctt	aacatggata	32400
cccctcttta	taccaagat	ggaaaattat	ccttacaagt	ttctccaccg	ttaaacadat	32460
taaaatcaac	cattctgaac	acattagctg	tagcttatgg	atcaggttta	ggactgagtg	32520
gtggcactgc	tcttgacagta	cagttggcct	ctccactcac	ttttgatgaa	aaaggaaata	32580
ttaaaattaa	cctagccagt	ggtccattaa	cagttgatgc	aagtcgactt	agtatcaact	32640
gcaaaagagg	ggtcactgtc	actacctcag	gagatgcaat	tgaaagcaac	ataagctggc	32700
ctaaaggat	aagatttgaa	ggtaatggca	tagctgcaaa	cattggcaga	ggatttggat	32760
ttggaaccac	tagtacagag	actgatgtca	cagatgcata	cccaattcaa	gttaaattgg	32820
gtactggcct	tacctttgac	agtacaggcg	ccattgttgc	ttggaacaaa	gaggatgata	32880
aacttacatt	atggaccaca	gccgaccct	cgccaaattg	caaaatatac	tctgaaaaag	32940
atgccaaact	cacactttgc	ttgacaaagt	gtggaagtca	aattctgggt	actgtgactg	33000
tattggcagt	gaataatgga	agtctcaacc	caatcacaaa	cacagtaagc	actgcactcg	33060
tctccctcaa	gtttgatgca	agtggagttt	tgctaagcag	ctccacatta	gacaaagaat	33120
attggaactt	cagaaaggga	gatgttacac	ctgctgagcc	ctataactaat	gctataggtt	33180

ttatgcctaa	cataaaggcc	tatcctaaaa	acacatctgc	agcttcaaaa	agccatattg	33240
tcagtcaagt	ttatctcaat	ggggatgagg	ccaaaccact	gatgctgatt	attactttta	33300
atgaaactga	ggatgcaact	tgcacctaca	gtatcacttt	tcaatggaaa	tgggatagta	33360
ctaagtacac	aggtgaaaca	cttgctacca	gctccttcac	cttctcctac	atcgcccaag	33420
aatgaacact	gtatcccacc	ctgcatgcc	acccttccca	ccccactctg	tctatggaaa	33480
aaactctgaa	gcacaaaata	aaataaagtt	caagtgtttt	attgattcaa	cagttttaca	33540
ggattcgagc	agttattttt	cctccaccct	cccaggacat	ggaatacacc	accctctccc	33600
cccgcacagc	cttgaacatc	tgaatgccat	tggtgatgga	catgcttttg	gtctccacgt	33660
tccacacagt	ttcagagcga	gccagtctcg	ggtcggtcag	ggagatgaaa	ccctccgggc	33720
actcccgcat	ctgcacctca	cagctcaaca	gctgaggatt	gtcctcggtg	gtcgggatca	33780
cggttatctg	gaagaagcag	aagagcggcg	gtgggaatca	tagtccgcga	acgggatcgg	33840
ccggtggtgt	cgcatcaggc	cccgcagcag	tcgctgccgc	cgccgctccg	tcaagctgct	33900
gctcaggggg	tccgggtcca	gggactccct	cagcatgatg	cccacggccc	tcagcatcag	33960
tcgtctggtg	cggcggggcg	agcagcgc	gcggatctcg	ctcaggctcg	tgcagtacgt	34020
gcaacacagg	accaccaggt	tgttcaacag	tccatagttc	aacacgctcc	agccgaaact	34080
catcgcgga	aggatgctac	ccacgtggcc	gtcgtaccag	atcctcaggt	aatcaagtg	34140
gcgctccctc	cagaacacgc	tgcccacgta	catgatctcc	ttgggcatgt	ggcggttcac	34200
cacctcccgg	taccacatca	ccctctggtt	gaacatgcag	ccccggatga	tcctgcggaa	34260
ccacagggcc	agcaccgccc	cgcccgccat	gcagcgaaga	gaccccggtt	cccggcaatg	34320
gcaatggagg	accaccgct	cgtaccctg	gatcatctgg	gagctgaaca	agtctatgtt	34380
ggcacagcac	aggcatatgc	tcatgcatct	cttcagcact	ctcagctcct	cgggggtcaa	34440
aaccatatcc	cagggcacgg	ggaactcttg	caggacagcg	aaccccgag	aacagggcaa	34500
tcctcgcaca	taacttacat	tgtgcatgga	cagggtatcg	caatcaggca	gcaccgggtg	34560
atcctccacc	agagaagcgc	gggtctcggt	ctcctcacag	cgtggtaagg	gggccggccg	34620
atacgggtga	tggcgggacg	cggtgatcg	tgttcgcgac	cgtgtcatga	tgcagttgct	34680
ttcggacatt	ttcgtacttg	ctgtagcaga	acctggtccg	ggcgctgcac	accgatcgcc	34740
ggcggcggtc	ccggcgcttg	gaacgctcgg	tgttgaaatt	gtaaaacagc	cactctctca	34800
gaccgtgcag	cagatctagg	gcctcaggag	tgatgaagat	cccatcatgc	ctgatagctc	34860
tgatcacatc	gaccaccgtg	gaatgggcca	gaccagcca	gatgatgcaa	ttttgttggg	34920
tttcggtgac	ggcgggggag	ggaagaacag	gaagaaccat	gattaacttt	taatccaaac	34980
ggtctcggag	cacttcaaaa	tgaaggctcg	ggagatggca	cctctcgccc	ccgctgtgtt	35040
ggtggaaaat	aacagccagg	tcaaaggatga	tacggttctc	gagatgttcc	acggtggctt	35100
ccagcaaagc	ctccacgcgc	acatccagaa	acaagacaat	agcgaaagcg	ggaggggttct	35160
ctaattcctc	aatcatcatg	ttacactcct	gcaccatccc	cagataattt	tcatttttcc	35220
agccttgaat	gattcgaact	agttcctgag	gtaaatccaa	gccagccatg	ataaagagct	35280

```

cgcgagagc gccctccacc ggcattctta agcacaccct cataattcca agatattctg 35340
ctcctgggtc acctgcagca gattgacaag cggaatatca aaatctctgc cgcgatccct 35400
aagctcctcc ctcagcaata actgtaagta ctctttcata tcctctccga aatttttagc 35460
cataggacca ccaggaataa gattagggca agccacagta cagataaacc gaagtcctcc 35520
ccagtgcgca ttgccaaatg caagactgct ataagcatgc tggctagacc cggatgatgc 35580
ttccagataa ctggacagaa aatcacccag gcaattttta agaaaatcaa caaaagaaaa 35640
atcctccagg tgcacgttta gagcctcggg aacaacgatg aagtaaagtc aagcgggtgcg 35700
ttccagcatg gttagttagc tgatctgtaa aaaacaaaaa ataaaacatt aaaccatgct 35760
agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggcca cgggggtctcc 35820
ggcgcgaccc tcgtaaaaaat tgtcgctatg attgaaaacc atcacagaga gacgttcccg 35880
gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg cgtccgcgag 35940
tgaaaaaaag cgcccgagga agcaataagg cactacaatg ctcagtctca agtccagcaa 36000
agcgatgcc a tgcggatgaa gcacaaaatc ctccaggtgc taaaaatgt aattactccc 36060
ctcctgcaca ggcagcgaag ccccgatcc ctccagatac acatacaaag cctcagcgtc 36120
catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
acctgtccac ccgctctctg ctcaatatat agcccagatc tacactgacg taaaggccaa 36240
agtctaaaaa taccgcgcaa ataatcacac acgcccagca cacgcccaga aaccggtgac 36300
acactcaaaa aaatacgcgc acttcctcaa acgcccacac tgccgtcatt tccgggttcc 36360
cacgctacgt catcggaatt cgactttcaa attccgtcga ccgttaaaaa cgtcacccgc 36420
cccggcccta acggtcgcgc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
cagctcattt gcatattaac gcgcaccaa agtttgaggt atattattga tgatg 36535

```

```

<210> 4
<211> 34264
<212> DNA
<213> simian adenovirus SV-1

```

```

<400> 4
tccttattct ggaaacgtgc caatatgata atgagcgggg aggagcgagg cggggccggg 60
gtgacgtgag gtgacgtggg gtgacgcggg gtggcgcgag ggcggggcgg gagtggggag 120
gcgcttagtt ttacgtatg cggaaggagg ttttataacc gaagttgggt aatttgggag 180
tatacttgta agttttgtgt aatttggcgc gaaaaccggg taatgaggaa gttgaggtta 240
atatgtactt tttatgactg ggcggaattt ctgctgatca gcagtgaact ttgggcgctg 300
acggggagggt ttcgctacgt ggcagtacca cgagaaggct caaagggtccc atttattgta 360
ctcctcagcg ttttcgctgg gtattttaaac gctgtcagat catcaagagg cactcttga 420
gtgccggcga gtagagtttt ctctcccgcg ctgccgcgat gaggctgggt cccgagatgt 480
acgggtgttt ctgcagcgag acggcccgga actcagatga gctgcttaat acagatctgc 540
tggtgtttcc caactcgcct gtggcttcgc ctccgtcgct tcatgatctt ttcgatgtgg 600

```

aagtggatcc accgcaagat cccaacgagg acgcggtaaa cagtatgttc cctgaatgtc	660
tgtttgaggc ggctgaggag ggttctcaca gcagtgaaga gagcagacgg ggagaggaac	720
tggacttgaa atgctacgag gaatgtctgc cttctagcga ttctgaaacg gaacagacag	780
ggggagacgg ctgtgagtcg gcaatgaaaa atgaacttgt attagactgt ccagaacatc	840
ctggatcatgg ctgccgtgcc tgtgcttttc atagaaatgc cagcggaaat cctgagactc	900
tatgtgctct gtgttatctg cgccttacca gcgattttgt atacagtaag taaagtgttt	960
tcattggcgt acggtagggg attcgttgaa gtgctttgtg acttattatg tgtcattatt	1020
tctaggtgac gtgtccgacg tggaagggga aggagataga tcaggggctg ctaattctcc	1080
ttgcactttg ggggctgtgg ttccagttgg cttttttaa ccgagtgggtg gaggagaacg	1140
agccggagga gaccgagaat ctgagagccg gcctggaccc tccagtggaa gactaggtgc	1200
tgaggatgat cctgaagagg ggactagtgg ggggtgctagg aaaaagcaaa aaactgagcc	1260
tgaacctaga aactttttga atgagttgac tgtaagccta atgaatcggc agcgtcctga	1320
gacggtgttt tggactgagt tggaggatga gttcaagaag ggggaattaa acctcttgta	1380
caagtatggg tttgagcagt tgaaaactca ctggttgag ccgtgggagg atatggaaat	1440
ggctctagac acctttgcta aagtggctct gcggccggat aaagtttaca ctattcgccg	1500
cactgttaat ataaaaaaga gtgtttatgt tatcggccat ggagctctgg tgcaggtgca	1560
gacccagac cgggtggctt tcaattgcgg catgcagagt ttgggccccg ggggtgatagg	1620
tttgaatgga gttacatttc aaaatgtcag gtttactggg gatgatttta atggctctgt	1680
gtttgtgact agcaccagc taaccctcca cgggtgtttac tttttaact ttaacaatac	1740
atgtgtggag tcatggggtg ggggtgtctt gaggggctgc agttttcatg gttgctggaa	1800
ggcggtggtg ggaagaatta aaagtgtcat gtctgtgaag aaatgcatat ttgaacgtg	1860
tgtgatagct ctacgagtag aggggtacgg acggatcagg aataacgccg catctgagaa	1920
tggatgtttt cttttgctga aaggtagggc cagcgtaaag cataatatga tttgcggcag	1980
cggcctgtgc ccctcgagc tcttaacttg cgcagatgga aactgtcaca ccttgccgac	2040
cgtgcacata gtgtccact cgcgccgac ctggccaaca tttgagcaca atatgctcat	2100
gcgttgccg gttcacctag gtgctagacg cggcgtgttt atgccttatc aatgtaactt	2160
tagtcatact aagattttgc tggaaactga ttccttcctt cgagtatgtt tcaatggggt	2220
gtttgacatg tcaatggaac tttttaaaagt gataagatat gatgaaacca agtctcgttg	2280
tcgctcatgt gaatgcggag ctaatcattt gaggttgat cctgtaacct tgaacgttac	2340
cgaggagctg aggacggacc accacatgct gtcttgccgt cgtaccgact atgaatccag	2400
cgatgaggag tgaggtgagg ggcggagcca caaaggggtat aaaggggcat gagggggtggg	2460
cgcggtgttt caaaatgagc gggacgacgg acggcaatgc gtttgagggg ggagtgttca	2520
gcccatatct gacatctcgt cttccttcct gggcaggagt tcgtcagaat gtagtgggct	2580
ccaccgtgga cggacggccg gtcgcccctg caaattccgc caccctcacc tatgccaccg	2640

tgggatcatc	gttggacact	gccgcggcag	ctgccgcttc	tgctgccgct	tctactgctc	2700
gcggcatggc	ggctgatttt	ggactatata	accaactggc	cactgcagct	gtggcgcttc	2760
ggtctctggt	tcaagaagat	gccctgaatg	tgatcttgac	tcgcctggag	atcatgtcac	2820
gtcgcctgga	cgaactggct	gcgcagatat	cccaagctaa	ccccgatacc	gcttcagaat	2880
cttaaaataa	agacaaacaa	atttgttgaa	aagtaaaatg	gctttatttg	tttttttttg	2940
ctcggtaggc	tcgggtccac	ctgtctcggt	cgtaaggac	tttgtgtatg	ttttccaaaa	3000
cacggtacag	atgggcttgg	atgttcaagt	acatgggcat	gaggccatct	ttgggggtgga	3060
gataggacca	ctgaagagcg	tcattgttccg	gggtgggtatt	gtaaatcacc	cagtcgtagc	3120
agggtttttg	agcgtggaac	tggaatatgt	ccttcaggag	caggctaata	gccaagggtg	3180
gacccttagt	gtaggtgttt	acaaagcggt	tgagctggga	gggatgcatg	cgggggggaga	3240
tgatatgcat	cttggcttgg	attttgaggt	tagctatgtt	accacccagg	tctctgcggg	3300
ggttcatgtt	atgaaggacc	accagcacgg	tatagccagt	gcatttgggg	aacttgtcat	3360
gcagtttgga	ggggaaggcg	tggaagaatt	tagatacccc	cttgtgcccc	cctaggtttt	3420
ccatgcactc	atccataata	atggcaatgg	gacccttggc	ggccgcttta	gcaaacacgt	3480
tttggggggtt	ggaaacatca	tagttttgct	ctagagttag	ctcatcatag	gccatcttta	3540
caaagcgggg	taggagggtg	cccgaactgg	ggatgatagt	tccatctggg	cctggagcgt	3600
agttgccctc	acagatctgc	atctcccagg	ccttaatttc	cgaggggggg	atcatgtcca	3660
cctggggggc	gataaaaaac	acggtttctg	gcgggggggtt	aatgagctgg	gtggaaagca	3720
agttacgcaa	cagctgggat	ttgccgcaac	cggtgggacc	gtagatgacc	ccgatgacgg	3780
gttgacagctg	gtagttcaga	gaggaacagc	tgccgtcggg	gcgcaggagg	ggagctacct	3840
cattcatcat	gcttctgaca	tgtttttttt	cactcactaa	gttttgcaag	agcctctccc	3900
caccagggga	taagagttct	tccaggctgt	tgaagtgttt	cagcggtttc	aggccgtcgg	3960
ccatgggcat	cttttcaagc	gactgacgaa	gcaagtacag	tcggtcccag	agctcgggtga	4020
cgtgctctat	ggaatctcga	tccagcagac	ttcttggttt	cggggggttg	gccgactttc	4080
gctgtagggc	accagccggt	gggcgtccag	ggccgcgagg	gttctgtcct	tccagggtct	4140
cagcgttcgg	gtgagggtg	tctcgggtgac	ggtgaaggga	tgagccccgg	gctgggcgct	4200
tgcgagggtg	cgcttcaggc	tcattcctgct	ggtgctgaag	cgggcgctcg	ctccctgtga	4260
gtcggccaga	tagcaacgaa	gcatgaggct	gtagctgagg	gactcggccg	cgtgtccctt	4320
ggcgcgagc	tttcccttgg	aaacgtgctg	acatttggtg	cagtgcagac	acttgagggc	4380
gtagagtttt	ggggccagga	agaccgactc	gggcgagtag	gcgtcggctc	cgactgagc	4440
gcagacggct	tcgcactcca	ccagccacgt	gagctcgggt	ttagcgggat	caaaaaccaa	4500
gttgccctcca	ttttttttga	tgcgtttctt	accttgctgc	tccatgagtc	tgtgtcccg	4560
ttccgtgaca	aaaaggctgt	cggtatcccc	gtagaccgac	ttgagggggc	gatcttccaa	4620
agggtgttccg	aggcttccg	cgtagaggaa	ctgggaccac	tccgagacaa	aggctcgggt	4680
ccaggctaac	acgaaggagg	cgatctgcga	ggggtatctg	tcgttttcaa	tgaggggggtc	4740

caccttttcc	aggggtgtgca	gacacaggtc	gtcctcctcc	gcgtccacga	aggtgattgg	4800
cttghtaagt	taggtcacgt	gacccgcacc	cccccaaggg	gtataaaagg	gggcgtgccc	4860
actctccccg	tcacttttctt	ccgcatcgct	gtggaccaga	gccagctgtt	cgggtgagta	4920
ggccctctca	aaagccggca	tgattttcggc	gctcaagttg	tcagtttcta	caaacgaggt	4980
ggatttgata	ttcacgtgcc	ccgcggcgat	gcttttgatg	gtggaggggt	ccatctgatc	5040
agaaaacacg	atcttttttat	tgtcaagttt	ggtggcgaaa	gacccgtaga	gggcgttga	5100
aagcaacttg	gcgatggagc	gcaggggtctg	atttttctcc	cgatcggccc	tctccttggc	5160
ggcgatgttg	agttgcacgt	actcgcgggc	cacgcaccgc	cactcgggga	acacggcggt	5220
gcgctcgtcg	ggcaggatgc	gcacgcgcca	gccgcggttg	tgcaggggtga	tgagggtccac	5280
gctggtggcc	acctccccgc	ggagggggctc	gttgggtccaa	cacaatcgcc	cccccttttct	5340
ggagcagaac	ggaggcaggg	gatctagcaa	gttggcgggc	ggggggtcgg	cgtcgatggt	5400
aaatatgccg	ggtagcagaa	ttttattaaa	ataatcgatt	tcggtgtccg	tgtcttgcaa	5460
cgcgctcttcc	cacttcttca	ccgccagggc	cttttcgtag	ggattcaggg	gcgggtcccca	5520
gggcatgggg	tgggtcaggg	ccgaggcgta	catgccgcag	atgtcgtaca	cgtacagggg	5580
ctccctcaac	accccgatgt	aagtggggta	acagcgcccc	ccgcggatgc	tggctcgcac	5640
gtagtcgtac	atctcgtgag	agggagccat	gagcccgtct	cccaagtggg	tcttgtgggg	5700
tttttcggcc	cggtagagga	tctgcctgaa	gatggcggtg	gagttggaag	agatagtggg	5760
gcgttggaag	acgttaaagt	tggctccggg	cagtccacg	gagtcttgga	tgaactgggc	5820
gtaggattcc	cggagcttgt	ccaccagggc	tgcggttacc	agcacgtcga	gagcgcagta	5880
gtccaacgtc	tcgcggacca	ggttgtaggc	cgtctcttgt	tttttctccc	acagttcgcg	5940
attgaggagg	tattcctcgc	ggtctttcca	gtactcttcg	gcgggaaatc	ctttttcgtc	6000
cgctcggtaa	gaacctaac	tgtaaaattc	gttcacggct	ttgtatggac	aacagccttt	6060
ttctaccggc	agggcgtagc	cttgagcggc	ctttctgaga	gaggtgtggg	tgagggcgaa	6120
ggtgtccccg	accatcactt	tcaggtactg	atgtttgaag	tccgtgtcgt	cgcaggcgcc	6180
ctgttcccac	agcgtgaagt	cgggtgcgctt	tttctgcctg	ggattgggga	gggcgaatgt	6240
gacgtcgtta	aagaggattt	tcccggcgcg	gggcatgaag	ttgcgagaga	tcctgaaggg	6300
tccgggcacg	tccgagcggg	tgttgatgac	ttgcgcggcc	aggacgatct	cgtcgaagcc	6360
gttgatgttg	tggcccacga	tgtaaagtcc	gataaagcgc	ggctgtccct	tgagggccgg	6420
cgcttttttc	aactcctcgt	aggtgagaca	gtccggcgag	gagagacca	gctccgcccg	6480
ggcccagtcg	gagagctgag	ggttagccgc	gaggaaagag	ctccacaggt	caagggctag	6540
cagagtttgc	aagcggtcgc	ggaactcgcg	aaactttttc	cccacggcca	ttttctccgg	6600
cgtcaccacg	tagaaagtgc	aggggcgggc	gttccagacg	tcccatcgga	gctctagggc	6660
cagctcgcag	gcttgacgaa	cgaggggtctc	ctcgcccgag	acgtgcatga	ccagcatgaa	6720
gggtaccaac	tgtttcccga	acgagcccat	ccatgtgtag	gtttctacgt	cgtaggtgac	6780

aaagagccgc	tgggtgcgcg	cgtgggagcc	gatcggaag	aagctgatct	cctgccacca	6840
gttgaggaa	tgggtgttga	tgtggtgaaa	gtagaagtcc	cgccggcgca	cagagcattc	6900
gtgctgatgt	ttgtaaaagc	gaccgcagta	gtcgcagcgc	tgacagctct	gtatctcctg	6960
aatgagatgc	gcttttcgcc	cgcgaccag	aaaccggagg	gggaagttga	gacgggggct	7020
tgggtgggcg	gcatccccctt	cgccttggcg	gtgggagtct	gcgtctgcgc	cctccttctc	7080
tgggtggacg	acggtgggga	cgacgacgcc	ccgggtgccg	caagtccaga	tctccgccac	7140
ggagggggcg	aggcgttgca	ggaggggacg	cagctgcccc	ctgtccaggg	agtcgagggc	7200
ggccgcgctg	aggtcggcgg	gaagcgtttg	caagttcact	ttcagaagac	cggtaaagagc	7260
gtgagccagg	tgacatggt	acttgatttc	caggggggtg	ttggaagagg	cgtccacggc	7320
gtagaggagg	ccgtgtccgc	gcggggccac	caccgtgccc	cgaggagggt	ttatctcact	7380
cgtcgagggc	gagcgccggg	gggtagaggc	ggctctgcgc	cggggggcag	cgagggcagt	7440
ggcacgtttt	cgtgaggatt	cggcagcggg	tgatgacgag	cccggagact	gctggcgtgg	7500
gcgacgacgc	ggcggttgag	gtcctggatg	tgccgtctct	gcgtgaagac	caccggcccc	7560
cgggtcctga	acctgaaaga	gagttccaca	gaatcaatgt	ctgcatcggt	aacggcgggc	7620
tgccctgagga	tctcctgtac	gtcgcccag	ttgtcttgat	aggcgatctc	ggccatgaac	7680
tgctccactt	cttcctcgcg	gaggtcgccg	tggcccgtc	gctccacggg	ggcgccagg	7740
tcgttgagga	tgcgacgcat	gagttgagag	aaggcggtga	ggccgttctc	gttccacacg	7800
cggctgtaca	ccacgtttcc	gaaggagtcg	cgcgctcgca	tgaccacctg	ggccacgttg	7860
agttccacgt	ggcgggcgaa	gacggcgtag	tttctgaggc	gctggaagag	gtagttgagc	7920
gtggtggcga	tgtgctcgca	gacgaagaag	tacatgatcc	agcgccgcag	ggcatctctg	7980
ttgatgtctc	cgatggcttc	gagacgctcc	atggcctcgt	agaagtcgac	ggcgaagttg	8040
aaaaattggg	agttgcgggc	ggccaccgtg	agttcttctt	gcaggaggcg	gatgagatcg	8100
gcgaccgtgt	cgcgcacctc	ctgctcgaaa	gcgccccgag	gcgcctctgc	ttcttcctcc	8160
ggctcctcct	cttcaggggg	cacgggttcc	tccggcagct	ctgcgacggg	gacggggcg	8220
cgacgtcgtc	gtctgaccgg	caggcggtcc	acgaagcgct	cgatcatttc	gccgcgccgg	8280
cgacgcatgg	tctcggtgac	ggcgcgctcc	tttctcgag	gtcgcagttc	gaagacgccg	8340
ccgcgcagag	cgcggcggtg	cagggagggg	aagtgggttag	ggccgtcggg	cagggacacg	8400
gcgctgacga	tgcattttat	caattgctgc	gtaggcactc	cgtgcaggga	tctgagaacg	8460
tcgaggtcga	cgggatccga	gaacttctct	aggaaagcgt	ctatccaatc	gcagtcgcaa	8520
ggtaagctga	ggacggtggg	ccgctggggg	gcgtccgcgg	gcagttggga	ggtgatgctg	8580
ctgatgatgt	aattaaagta	ggcggtcttc	aggcggcgga	tggtggcgag	gaggaccacg	8640
tctttgggcc	cggcctgttg	aatgcgcagg	cgctcggcc	tgccccaggc	ctcgctctga	8700
cagcgacgca	ggtcttttga	gtagtcttgc	atcagtctct	ccaccggaac	ctctgcttct	8760
cccctgtctg	ccatgcgagt	cgagccgaac	ccccgcaggg	gctgcagcaa	cgctaggtcg	8820
gccacgaccc	tctcggccag	cacggcctgt	tggatctgcg	tgagggtggg	ctggaagtcg	8880

tccaggtcca	cgaagcggtg	ataggccccc	gtgttgatgg	tgtaggtgca	gttggccatg	8940
acggaccagt	tgacgacttg	catgccgggt	tgggtgatct	ccgtgtactt	gaggcgcgag	9000
taggcgcggg	actcgaacac	gtagtcgttg	catgtgcgta	ccagatactg	gtagccaacc	9060
aggaagtggg	gaggcggttc	tcggtacagg	ggccagccga	ctgtggcggg	ggcgccgggg	9120
gacaggtcgt	ccagcatgag	gcgatggtag	tggtagatgt	agcgggagag	ccaggtgatg	9180
ccggccgagg	tggtcgcggc	cctggtgaat	tcgcggacgc	ggttccagat	gttgcgcagg	9240
gggcgaaagc	gctccatggt	gggcacgctc	tgccccgtga	ggcgggcgca	atcttgtagc	9300
ctctagatgg	aaaaaagaca	gggcgggtcat	cgactccctt	ccgtagctcg	gggggtaaag	9360
tcgcaagggg	gcggcgggcg	ggaaccccg	ttcgagaccg	gccggatccg	ccgctcccga	9420
tgcgccctggc	cccgcattcca	cgacgtccgc	gtcgagaccc	agccgcgacg	ctccgccccca	9480
atacggaggg	gagtcttttg	gtgttttttc	gtagatgcat	ccggtgctgc	ggcagatgcg	9540
acctcagacg	cccaccacca	ccgccgcggc	ggcagtaaac	ctgagcggag	gcggtgacag	9600
ggaggaggag	gagctggctt	tagacctgga	agagggagag	gggctggccc	ggctgggagc	9660
gccgtcccca	gagagacacc	ctagggttca	gctcgtgagg	gacgccaggc	aggcttttgt	9720
gccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	tgcgcgattg	9780
caggttttcg	gcgggtagag	agctgagggc	gggcttcgat	cgggagcggc	tcctgagggc	9840
ggaggatttc	gagcccagcg	agcgttctgg	ggtgagccc	gcccgcgctc	acgtctcggc	9900
ggccaacctg	gtgagcgcg	acgagcagac	ggtgaacgag	gagcgcaact	tccaaaagag	9960
ctttaacaat	cacgtgagga	ccctgatcgc	gagggaggag	gtgaccatcg	ggctgatgca	10020
tctgtgggac	ttcgtggagg	cctacgtgca	gaacccggcc	agcaaacctc	tgacggccca	10080
gctgttcctg	atcgtgcagc	acagccgcga	caacgagacg	ttccgcgacg	ccatgttgaa	10140
catcgcgagg	cccaggggtc	gctggctctt	ggatctgatt	aacatcctgc	agagcatcgt	10200
ggtgcaggag	aggggcctca	gcttagcgga	caagggtggc	gccattaact	attcgatgca	10260
gagcctgggg	aagtctctac	ctcgcaagat	ctacaagagc	ccttacgtgc	ccatagacaa	10320
ggaggtgaag	atagacagct	tttacctg	catggcgctg	aaggtgctga	cgctgagcga	10380
cgacctcggc	gtgtaccgta	acgacaagat	ccacaaggcg	gtgagcgcca	gccgccggcg	10440
ggagctgagc	gacagggagc	tgatgcacag	cctgcagagg	gcgctggcgg	gcgccgggga	10500
cgaggagcgc	gaggcttact	tcgacatggg	agccgatctg	cagtggcgctc	ccagcgcgcg	10560
cgctttggag	gcggcgggct	accccgacga	ggaggatcgg	gacgatttgg	aggaggcagg	10620
cgagtacgag	gacgaagcct	gaccgggcag	gtgttgtttt	agatgcagcg	gccggcgggac	10680
ggggccaccg	cggatcccg	acttttggca	tccatgcaga	gtcaaccttc	gggcgtgacc	10740
gcctccgatg	actgggcggc	ggccatggac	cgcatatg	cgctgactac	ccgcaacccc	10800
gaggctttta	gacagcaacc	ccaggccaac	cgtttttcgg	ccatcttgga	agcgggtggtg	10860
ccctcccgca	ccaacccac	acacgagaaa	gtcctgacta	tcgtgaacgc	cctggtagac	10920

agcaaggcca	tccgccgcga	cgaggcgggc	ttgatttaca	acgctctgct	ggaacgggtg	10980
gcgcgctaca	acagcactaa	cgttcagacc	aatctggatc	gcctcaccac	cgacgtgaag	11040
gaggcgctgg	ctcagaagga	gcggtttctg	agggacagca	atctgggctc	tctggtggca	11100
ctcaacgcct	tcctgagcac	gcagccggcc	aacgtgcccc	gcgggcagga	ggactacgtg	11160
agcttcatca	gcgctctgag	gctgctggtg	tccgaggtgc	cccagagcga	ggtgtatcag	11220
tctgggccgg	attacttctt	ccagacgtcc	cgacagggct	tgcaaacggt	gaacctgact	11280
caggccttta	aaaacttgca	aggcatgtgg	ggcgtaag	ccccgggtgg	cgatcgagcc	11340
accatctcca	gtctgctgac	ccccaacact	cgcttgcgtc	tgctcttgat	cgcgccgttc	11400
accaacagta	gcactatcag	ccgtgactcg	tacctgggtc	atctcatcac	tttgtaccgc	11460
gaggccatcg	gtcaggctca	gatcgacgag	cacacatatc	aggagatcac	taacgtgagc	11520
cgggccctgg	gtcaggaaga	taccggcagc	ctggaagcca	cgttgaactt	tttgctaacc	11580
aaccggaggc	aaaaaatacc	ctcccagttt	acgttaagcg	ccgaggagga	gaggattctg	11640
cgatacgtgc	agcagtcctg	gagtcctgtac	ttgatgcggg	agggcgccac	cgcttccacg	11700
gctttagaca	tgacggctcg	gaacatggaa	ccgtcctttt	actccgcccc	ccggccgttc	11760
attaaccgtc	tgatggacta	cttccatcgc	gcggccgcca	tgaacgggga	gtacttcacc	11820
aatgccatcc	tgaatccgca	ttggatgccc	ccgtccggct	tctacaccgg	cgagtttgac	11880
ctgcccgaag	ccgacgacgg	ctttctttgg	gacgacgtgt	ccgacagcat	tttcacgccg	11940
ggcaatcgcc	gattccagaa	gaaggagggc	ggagacgagc	tccccctctc	cagcgtggag	12000
gcggcctcta	ggggagagag	tccctttccc	agtctgtctt	ccgccagcag	tggtcgggta	12060
acgcgcccgc	ggttgccggg	ggagagcgac	tacctgaacg	accccttgct	gcggccggct	12120
aggaagaaaa	atttcccaa	caacgggggtg	gaaagcttgg	tggataaaat	gaatcgttgg	12180
aagacctacg	cccaggagca	gcgggagtg	gaggacagtc	agccgcgacc	gctggttccg	12240
ccgcactggc	gtcgtcagag	agaagacccg	gacgactccg	cagacgatag	tagcgtgttg	12300
gacctgggag	ggagcggagc	caaccctttt	gctcacttgc	aaccaagg	gcgttccagt	12360
cgctctact	aataaaaaag	acgcggaaac	ttaccagagc	catggccaca	gcgtgtgtcc	12420
tttcttcctc	tctttcttcc	tcggcgcggc	agaatgagaa	gagcgggtgag	agtcacgccg	12480
gcggcgatatg	agggtccgcc	cccttcttac	gaaagcgtga	tgggatcagc	gaacgtgccg	12540
gccacgtgg	aggcgctta	cgttcctccc	agatacctgg	gacctacgga	gggcagaaac	12600
agcatccgtt	actccgagct	ggcacccttg	tacgatacca	ccaagggtga	cctggtggac	12660
aacaagtcgg	cggacatcgc	ctccctgaat	tatcaaaacg	atcacagcaa	ttttctgact	12720
accgtggtgc	agaacaatga	cttcaccccc	acggaggcgg	gcacgcagac	cattaacttt	12780
gacgagcgtt	cccgtgggg	cggtcagctg	aaaaccatcc	tgacaccaa	catgccaac	12840
atcaacgagt	tcagtgtccac	caacaagtgc	agggccaggc	tgatggttaa	aaaggctgaa	12900
aaccagcctc	ccgagtacga	atggtttgag	ttcaccattc	ccgagggcaa	ctattccgag	12960
accatgacta	tcgatctgat	gaacaatgcg	atcggtggaca	attacctgca	agtggggagg	13020

cagaacgggg tattggaaag cgatatcggc gtaaaatttg ataccagaaa cttccgactg 13080
gggtgggatc ccgtgaccaa gctggtgatg ccaggcgtgt acaccaacga ggcttttcac 13140
cccgacatcg tgctgctgcc ggggtgcggt gtggacttca ctgagagccg tttgagtaac 13200
ctgttaggga tcagaaagcg ccgccccttc caagagggtt ttcagatcat gtatgaggac 13260
ctggaaggag gtaacattcc aggtttgcta gacgtgccgg cgtatgaaga gagtggtaaa 13320
caggcggagg cgcaggggacg agagattcga ggcgacacct ttgccacgga acctcacgaa 13380
ctggaataaa aacctctgga acaagacagt aaaaaacgga gttacaacat tatatccggc 13440
actatgaata ccttgtaccg gagctggttt ctggcttaca actacgggga tcccgaag 13500
ggagttagat catggaccat actcaccacc acggacgtga cctgcggctc gcagcaagtg 13560
tactggtccc tgccggatat gatgcaagac ccggtcacct tccgcccctc caccgaagtc 13620
agcaacttcc cgggtggtggg caccgagctg ctgcccgtcc atgccaagag cttctacaac 13680
gaacaggccg tctactcgca actcattcgc cagtccaccg cgcttaccga cgtgttcaat 13740
cgctttcccg agaaccagat tctggtgctc cctcccgtc ctaccattac caccgtcagt 13800
gaaaacgttc ccgccctcac agatcacgga accctgccgc tgcgcagcag tatcagtga 13860
gttcagcgcg tgaccatcac cgacgccaga cgtcgaacct gtccctacgt ttacaaagct 13920
cttggcgtag tggctcctaa agtgctctct agtcgcacct tctaaacatg tccatcctca 13980
tctctcccga taacaacacc ggctggggac tgggctccgg caagatgtac ggcgagacca 14040
aaaggcgctc cagtcagcac ccagttcgag ttcggggcca cttccgtgct ccctggggag 14100
cttacaagcg aggactctcg ggccgaacgg cggtagacga taccatagat gccgtgattg 14160
ccgacgcccg ccggtacaac cccggaccgg tcgctagcgc cgctccacc gtggattccg 14220
tgatcgacag cgtggttagct ggcgctcggg cctatgctcg ccgcaagagg cggctgcac 14280
ggagacgtcg cccaccgcc gccatgctgg cagccagggc cgtgctgagg cgggcccgg 14340
gggtaggcag aagggtatg cgccgcgctg ccgccaacgc cgccgccggg agggcccgc 14400
gacaggctgc ccgccaggct gctgccgcca tcgctagcat ggccagacc aggagaggga 14460
acgtgtactg ggtgctgat tctgtgacgg gaggccgagt gccggtgctc agccgacctc 14520
cccgaagtta gaagatcaa gctgcgaaga cggcggtact gagtctccct gttgttatca 14580
gccaacatg agcaagcgca agtttaaaga agaactgctg cagacgctgg tgcctgagat 14640
ctatggccct ccggacgtga agcctgacat taagccccgc gatatcaagc gtgttaaaaa 14700
gcgggaaaag aaagaggaa tcgcggtggt agacgatggc ggagtgaat ttattaggag 14760
tttcgccccg cgacgcaggg ttcaatggaa agggcgggcg gtacaacgcg ttttgaggcc 14820
gggcaccgcg gtagttttta cccggggaga gcggtcggcc gttaggggtt tcaaaaggca 14880
gtacgacgag gtgtacggcg acgaggacat attggaacag gcggtcaac agatcggaga 14940
atgtgcctac ggaaagcgtt cgcgtcgcga agacctggcc atcgctttag acagcgga 15000
ccccacgccc agcctcaaac ctgtgacgct gcagcaggtg ctccccgtga gcgccagcac 15060

ggacagcaag aggggaataa aaagagaaat ggaagatctg cagcccacca tccagctcat 15120
 ggtccctaaa cggcagagggc tggaagaggt cctggagaaa atgaaagtgg acccaagcat 15180
 agagccggac gtcaaagtca ggccgatcaa agaagtggcc cctggtctcg ggggtgcagac 15240
 ggtggatatc cagatccccg tcacgtcagc ttcgaccgcc gtggaagcca tggaaacgca 15300
 aacggaaacc cctgccgcga tcggtaccag ggaagtggcg ttgcaaaccg acccctggta 15360
 cgaatacgcc gcccctcggc gtcagaggcg acccgctcgt tacggccccg ccaacgccat 15420
 catgccagaa tatgcgctgc atccgtctat cctgcccacc cccggctacc ggggagtgc 15480
 gtatcgcccc tcaggaaccc gccgccgaac ccgtcgccgc cgccgctccc gtcgtgctct 15540
 ggcccccgctg tcggtgcgcc gcgtaacacg ccggggaaag acagttacca ttcccaaccc 15600
 gcgctaccac cctagcatcc tttaatgact ctgccgtttt gcagatggct ctgacttgcc 15660
 gcgtgcgcct tcccgttccg cactatcgag gaagatctcg tcgtaggaga ggcatggcgg 15720
 gtagtggtcg ccggcgggct ttgcgcaggc gcatgaaagg cggaatttta cccgctctga 15780
 tacccataat cgccgccgcc atcggtgcca taccggcgt cgcttcagtg gccttgcaag 15840
 cagctcgtaa taaataaacg aaggcttttg cacttatgtc ctggtcctga ctattttatg 15900
 cagaaagagc atggaagaca tcaattttac gtcgctggct ccgcggcacg gctcgcggcc 15960
 gctcatgggc acctggaacg acatcggcac cagtcagctc aacgggggcg ctttcaattg 16020
 ggggagcctt tggagcggca ttaaaaactt tggctccacg attaaatcct acggcagcaa 16080
 agcctggaac agtagtgctg gtcagatgct ccgagataaa ctgaaggaca ccaacttcca 16140
 agaaaaagtg gtcaatgggg tggtagccgg catccacggc gcggtagatc tcgccaacca 16200
 agcggtgagc aaagagattg acaggcgttt ggaaagctcg cgggtgccgc cgcagagagg 16260
 ggatgaggtg gaggtcgagg aagtagaagt agaggaaaag ctgccccgcg tggagaaagt 16320
 tcccgggtcg cctccgagac cgcagaagcg acccaggcca gaactagaag aaactctggt 16380
 gacggagagc aaggagcctc cctcgtacga gcaagccttg aaagagggcg cctctccacc 16440
 ctaccaatg acaaaaccga tcgcgcttat ggctcggccg gtgtacggga aggactacaa 16500
 gcctgtcacg ctagagctcc ccccgccgcc accgccgcc cccacgcgcc cgaccgttcc 16560
 cccccccctg ccggctccgt cggcgggacc cgtgtccgca cccgtcgccg tgcctctgcc 16620
 agccgcccgc ccagtggccg tggccactgc cagaaacccc agaggccaga gaggagccaa 16680
 ctggcaaagc acgctgaaca gcatcgtggg cctgggagtg aaaagcctga aacgccgcg 16740
 ttgctattat taaaagtgtg gctaaaaaat ttcccgttgt atacgcctcc tatgttaccg 16800
 ccagagacgc gtgactgtcg ccgcgagcgc cgctttcaag atggccaccc catcgatgat 16860
 gccgcagtgg tcttacatgc acatcgccgg gcaggacgcc tcggagtacc tgagccccgg 16920
 tctcgtgcag ttcgcccgcg ccaccgacac ctacttcagc ttgggaaaca agtttagaaa 16980
 cccaccgtg gccccaccc acgatgtaac cacggaccgc tcgcaaaggc tgaccctgcg 17040
 ttttgtgccc gtagaccggg aggacaccgc gtactcttac aaagtgcgct acacgctggc 17100
 cgtaggggac aaccgagtgc tggacatggc cagcacctac tttgacatcc ggggagtgc 17160

ggatcgcggt cccagtttta agccctactc ggggtaccgcg tacaattccc tggctcccaa 17220
 gggcgctccc aaccctgcag aatggacgaa ttcagacagc aaagttaaag tgagggcaca 17280
 ggcgcctttt gttagctcgt atgggtgctac agcgattaca aaagagggta ttcaggtggg 17340
 agtaacctta acagactccg gatcaacacc acagtatgca gataaaacgt atcagcctga 17400
 gccgcaaatt ggagaactac agtggaaacag cgatgttgga accgatgaca aaatagcagg 17460
 aagagtgcta aagaaaacaa cgcccatgtt cccttggttac ggctcatatg ccaggcccac 17520
 taatgaaaaa ggaggacagg caacaccgtc cgctagtcaa gacgtgcaa atcccgaatt 17580
 acaatttttt gcctctacta atgtcgccaa tacacaaaaa gcagttctat atgcggagga 17640
 cgtgtcaatt gaagcgccag acactcactt ggtgttcaaa ccaacagtca ctgaaggcat 17700
 tacaagttca gaggtcttac tgaccaaca agctgctccc aaccgtcaa actacatagc 17760
 ctttagagat aattttattg gtctcatgta ctacaatagc acaggtaaca tgggagtact 17820
 ggcaggccag gcttctcagc taaatgcagt tgttgacctg caagacagaa atactgagct 17880
 gtcctaccaa ctcatgttg acgccctcgg agaccgcagt cggactttt ctatgtggaa 17940
 ccaagctgtg gatagttacg atcctgatgt aagaatcata gaaaaccatg gcgtagaaga 18000
 tgaattgcct aattattgct ttcctttggg aggcattggca gtaaccgaca cctactcgcc 18060
 tataaagggt aatggaggag gcaatggatg ggaagccaat aacggcggtt tcaccgaaag 18120
 aggagtggaa ataggttcag ggaacatgtt tgccatggag attaacctgc aagccaacct 18180
 atggcgtagc tttctgtact ccaatattgg gctgtacctg ccagactctc tcaaaatcac 18240
 tcctgacaac atcacactcc cagagaacaa aaacacctat cagtatatga acggtcgcgt 18300
 gacgccacc gggctgggtg acacctacgt taacgtgggc gcgcgctggg ccccgatgt 18360
 catggacagt attaacctt ttaatcacca ccgcaacgcc ggactccgct accgttccat 18420
 gctcctggga aacggacgct acgtgccctt ccacatccag gtgccccaga aattctttgc 18480
 aattaaaaac ctgctgctgc tccccggtt ctacacctac gagtggaaact tccgcaagga 18540
 cgtgaacatg atcttgca gctcgctggg caatgacctg cgagtggacg gggccagcat 18600
 ccgcttcgac agcatcaacc tgtacgcaa ctttttcccc atggcccaca acacggcctc 18660
 caccctggaa gccatgctgc gcaacgacac caacgacaa tctttcaacg actacctgtg 18720
 cgcgccaac atgctgtacc ccatccccgc caacgccacc agcgtgcca tctccattcc 18780
 ctctcgcaac tgggcagcct tcaggggctg gagtttcacc cgcctcaaaa ccaaggagac 18840
 cccctcgctg ggctccgggt tcgacccta cttcgtctac tccggctcca tcccctacct 18900
 ggacggcacc ttctacctca accatacttt caaaaagggtg tcaatcatgt tcgactcctc 18960
 cgtcagctgg cccggcaacg accgtctgct gacgccaac gagttcgaaa tcaagcgttc 19020
 ggtggacggt gaagggtaca acgtggctca gagcaacatg accaaggact ggttcctgat 19080
 tcagatgctc agccactaca acatcggcta ccagggcttc tacgtgcccg aaaattacaa 19140
 ggaccgcatg tactctttct tcagaaactt ccaacccatg agccgcaaaa ttgtagattc 19200

aacggccttac actaattatc aggatgtgaa actgccatac cagcataaca actcaggggtt 19260
 cgtgggcttac atgggaccca ccatgcgaga ggggcaggcc taccgggcca actatcccta 19320
 tcccctgatt ggggccaccg ccgtgcccag cctcacgcag aaaaagttcc tctgcgaccg 19380
 ggtgatgtgg aggatcccct tctctagcaa cttcatgtct atgggctccc tcaccgacct 19440
 ggggcagaac atgctgtacg ccaactccgc tcacgccttg gatatgacct ttgaggtgga 19500
 tcccatggat gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tgggtgcgcat 19560
 ccaccagccg caccgcgggc tcatcgaggc cgtctacctg cgcacacctt tctctgccgg 19620
 taacgccacc acctaaagaa gccgatgggc tccagcgaac aggagctgca ggccattgtt 19680
 cgcgacctgg gctgcgggcc ctactttttg ggcaccttcg acaagcgttt tcccggcttc 19740
 atgtccccc acaagccggc ctgtgccatc gttaacacgg ccggacggga gaccgggggg 19800
 gtccactggc tcgccttcgc ctggaacccg cgtaaccgca cctgctacct gttcgaccct 19860
 tttggtttct ccgacgaaag gctgaagcag atctaccagt tcgagtacga ggggctcctc 19920
 aagcgacgcg ctctggcctc cagccccgac cactgcgta ccctggaaaa gtccacccaa 19980
 acggtccagg ggcccccttc ggccgcctgc gggctcttct gttgcatgtt tttgcacgcc 20040
 ttcgtgcact ggcctcacac ccccatggat cacaaccca ccatggatct gctcaccgga 20100
 gtgccaaca gcatgcttca cagccccag gtcgccccca ccctgcgccg taaccaggaa 20160
 cacctgtatc gctttctggg gaaacactct gcctattttc gccgccaccg gcagcgcac 20220
 gaacgggcca cggccttcga aagcatgagc caaagagtgt aatcaataaa aaacattttt 20280
 atttgacatg atacgcgctt ctggcgtttt attaaaaatc gaagggttcg agggaggggt 20340
 cctcgtgccc gctggggagg gacacgttgc gatactggaa acgggcgctc caacgaaact 20400
 cggggatcac cagccgcggc aggggcacgt cttctagggt ctgcttccaa aactgccgca 20460
 ccagctgcag ggctcccatg acgtcgggcg ccgatatctt gaagtcgcag ttagggccgg 20520
 agctcccgcg gctgttgccg aacacgggggt tggcacactg gaacaccagc acgccgggggt 20580
 tgtggatact ggccagggcc gtcgggtcgg tcacctccga cgcattcaga tcctcggcgt 20640
 tgctcagggc aaacggggtc agcttgaca tctgccgcc aatctgggggt actaggtcgc 20700
 gcttgttgag gcagtcgcag cgcagagga tcaggatgcg tcgctgcccg cgttgcatga 20760
 tagggtaact cgccgccagg aactcctcca tttgacggaa ggccatctgg gctttgccgc 20820
 cctcgggtga gaatagcccg caggacttgc tagagaatac gttatgaccg cagttgacgt 20880
 cctccgcgca gcagcgggcg tcttcgttct tcagctgaac cacgttgccg cccaacgggt 20940
 tctggaccac cttggctcta gtgggggtgct ccttcagcgc ccgctgtccg ttctcgtgg 21000
 ttacatccat ttccaacacg tgctccttgc agaccatctc cactccgtgg aagcaaaaca 21060
 ggacgccctc ctgctgggta ctgcgatgct ccatacggc gcatccggtg ggctcccagc 21120
 tcttggtgtt taccgccgcg taggcttcca tgtaagccat aaggaatctg cccatcagct 21180
 cggatgaagg cttctgggtg gtgaagggtta gcggcaggcc gcggtgctcc tcgttcaacc 21240
 aagtttgaca gatcttgccg tacaccgctc cctgggtcggg cagaaactta aaagccgctc 21300

tgctgtcgtt	gtctacgtgg	aactttctcca	ttaacatcat	catggtttcc	atacccttct	21360
cccacgctgt	caccagtgg	ttgctgtcgg	ggttcttcac	caacacggcg	gtagaggggc	21420
cctcgccggc	cccgacgtcc	ttcatggtca	ttctttgaaa	ctccacggag	ccgtccgcgc	21480
gacgtactct	gcgacccgga	gggtagctga	agcccacctc	caccacgggtg	ccttcgccct	21540
cgctgtcggg	gacaatctcc	ggggatggcg	gcggcgcggg	tgtcgccttg	cgagccttct	21600
tcttgggagg	gagctgaggc	gcctcctgct	cgcgctcggg	gctcatctcc	cgcaagtagg	21660
gggtaatgga	gctgcctgct	tggttctgac	ggttggccat	tgtatcctag	gcagaaagac	21720
atggagctta	tgcgcgagga	aactttaacc	gccccgtccc	ccgtcagcga	cgaagatgtc	21780
atcgtcgaac	aggacccggg	ctacgttacg	ccgcccagg	atctggaggg	gcctgaccgg	21840
cgcgacgcta	gtgagcggca	ggaaaatgag	aaagaggagg	cctgctacct	cctggaaggc	21900
gacgttttgc	taaagcattt	cgccaggcag	agcaccatag	ttaaggaggc	cttgcaagac	21960
cgctccgagg	tgcccttggg	cgtcgccgcg	ctctcccagg	cctacgaggc	gaaccttttc	22020
tcgcctcgag	tgccctcgaa	gagacagccc	aacggcacct	gcgagcccaa	cccgcgactc	22080
aactttctacc	ccgtgttcgc	cgtaccagag	gcgctggcca	cctatcacat	ttttttcaaa	22140
aaccaacgca	tccccctatc	gtgccggggc	aaccgcaccg	cggccgatag	gaatctcagg	22200
cttaaaaacg	gagccaacat	acctgatatc	acgtcgctgg	aggaagtgcc	caagattttc	22260
gagggctctg	gtcgagatga	gaagcggggc	gcgaacgctc	tgcaaaaaga	acagaaagag	22320
agtcagaacg	tgctggtgga	gctggagggg	gacaacgcgc	gtctggccgt	cctcaaacgc	22380
tgcatagaag	tctcccactt	cgcctacccc	gccctcaact	tgccacccaa	agttatgaaa	22440
tcggtcatgg	atcagctgct	catcaagaga	gctgagcccc	tggatcccga	ccaccccgag	22500
gcggaaaact	cagaggacgg	aaagcccgtc	gtcagcgacg	aggagctcga	gcggtggctg	22560
gaaaccaggg	accccccaaca	gttgcaagag	aggcgcaaga	tgatgatggc	ggccgtgctg	22620
gtcaccgtgg	agctggaatg	cctgcaacgg	tttttcagcg	acgtggagac	gctacgcaaa	22680
atcggggaat	ccctgcacta	caccttcgc	cagggctacg	tccgccaggc	ctgcaagatc	22740
tccaacgtgg	agctcagcaa	cctggtctcc	tacatgggca	tcctccacga	gaaccggctg	22800
gggcagagcg	tgctgcactg	caccttgcaa	ggcgaggcgc	ggcgggacta	cgtgcgagac	22860
tgcatctacc	tcttcctcac	cctcacctgg	cagaccgcca	tgggcgtctg	gcagcagtgc	22920
ttggaagaga	gaaacctcaa	agagctagac	aaactcctct	gccgccagcg	gcgcgccttg	22980
tggtccggtt	tcagcgagcg	cacggtcgcc	agcgtcttg	cggacatcat	cttcccggag	23040
cgcctgatga	aaaccttgca	aaacggcctg	ccggatttca	tcagtcaaag	cattttgcaa	23100
aacttccgct	cttttgtcct	ggaacgctcc	gggatcttgc	ccgccatgag	ctgcgcgcta	23160
ccttctgact	ttgtccccct	ctcctaccgc	gagtgccttc	ccccactgtg	gagccactgc	23220
tacctcttcc	aactggccaa	ctttctggcc	taccactccg	acctcatgga	agacgtaagc	23280
ggagaggggt	tactggagtg	ccactgccgc	tgcaacctgt	gcacccccca	cagatcgctg	23340

gcctgcaaca ccgagctact cagcgaaacc caggtcatag gtaccttcga gatccagggg 23400
 cccagcagc aagaggggtgc ttccggcttg aagctcactc cggcgctgtg gacctcggct 23460
 tacttacgca aatttgtagc cgaggactac cacgcccaca aaattcagtt ttacgaagac 23520
 caatctcgac caccgaaagc cccctcacg gcctgctgca tcaccagag caagatcctg 23580
 gccaattgc aatccatcaa ccaagcgcg cgcgatttcc ttttgaaaaa gggtcggggg 23640
 gtgtacctgg acccccagac cggcgaggaa ctcaaccctg ccacactctc cgtcgaagca 23700
 gccccccga gacatgccgc ccaagggaaac cgccaagcag ctgatcgctc ggcagagagc 23760
 gaagaagcaa gagctgctcc agcagcaggt ggaggacgag gaagagatgt gggacagcca 23820
 ggcagaggag gtgtcagagg acgaggagga gatggaaagc tgggacagcc tagacgagga 23880
 ggaggacgag ctttcagagg aagaggcgac cgaagaaaaa ccacctgcat ccagcgcgcc 23940
 ttctctgagc cgacagccga agccccggcc cccgacgccc ccggccggct cactcaaagc 24000
 cagccgtagg tgggacgcca ccgaatctcc agcggcagcg gcaacggcag cgggtaaggc 24060
 caaacgcgag cggcggggggt attgctcctg gcgggcccac aaaagcagta ttgtgaactg 24120
 cttgcaacac tgcgggggaa acatctcctt tgcccagcgc tacctcctct tccatcacgg 24180
 tgtggccttc cctcgaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24240
 gctcgagaa aaaagctaag gcctcctccg ccgcgaggaa aaactccgcc gccgctgccg 24300
 ccgccaagga tccaccggcc accgaagagc tgagaaagcg catctttccc actctgtatg 24360
 ctatctttca gcaaagccgc gggcagcacc ctgagcgcg actgaaaata aaaaaccgct 24420
 ccttcgctc gctcaccgc agctgtctgt accacaagag agaagaccag ctgcagcgca 24480
 ccctggacga cgccgaagca ctgttcagca aatactgctc agcgtctctt aaagactaaa 24540
 agaccgcgc tttttcccc tcggccgcca aaaccacgt catcgccagc atgagcaagg 24600
 agattccac ccctacatg tggagctatc agcccagat gggcctggcc gcggggggcg 24660
 cccaggacta ctccagcaag atgaactggc tcagcgccgg cccccacatg atctcacgag 24720
 ttaacggcat ccgagcccac cgaaaccaga ttctcttaga acaggcgga atcaccgcca 24780
 caccgcggcg ccaactcaac ccgcctagtt ggcccgccgc ccaggtgtat caggaaaatc 24840
 cccgcccagc cacagtcctc ctgccacgc acgcgagggc cgaagtcctc atgactaact 24900
 ctggggtaca attagcgggc gggtcaggt acgccaggta cagaggtcgg gccgctcctt 24960
 actctcccg gagtataaag agggatgatc ttcgaggccg aggtatccag ctcaacgacg 25020
 agacggtgag ctctcaacc ggtctcagac ctgacggagt cttccagctc ggaggagcgg 25080
 gccgctcttc cttcaccact cgccaggcct acctgaccct gcagagctct tcctcgagc 25140
 cgcgctccgg gggaatcggc actctccagt ttcgtggaaga gttcgttccc tccgtctact 25200
 tcaaccctt ctccggctcg cctggacgct accggagcgc cttcattccc aactttgacg 25260
 cagtgagtga atccgtggac ggctacgact gatgacagat ggtgcggccg tgagagctcg 25320
 gctgcgacat ctgcatcact gccgtcagcc tcgctgctac gctcgggagg cgatcgtctt 25380
 cagctacttt gagctgccgg acgagcacc tcagggtccg gctcacgggt tgaaactcga 25440

gatcgcgagaac gcgcctcgagt ctcgcctcat cgacaccttc accgcccgcac ctctcctggt 25500
agaaatccaa cgggggatca ctaccatcac cctgttctgc atctgccccca cgcccggatt 25560
acatgaagat ctgtgttggtc atctttgcgc tcagtttaaat aaaaactgaa ctttttgccg 25620
caccttcaac gccatctgtg atttctacaa caaaaagttc ttctggcaaa ggtacacaaa 25680
ctgtatttta ttctaattct acctcatcta tcgtgctgaa ctgcgcctgc actaacgaac 25740
ttatccagtg gattgcaaac ggtagtgtgt gcaagtactt ttgggggaac gatatagtta 25800
gtagaaataa cagcctttgc gagcactgca actcctccac actaatcctt tatcccccat 25860
ttgttactgg atggatatatg tgcgttggct ccggtttaaa tcctagtgtc tttcataagt 25920
ggtttctaca aaaagagacc cttcccaaca attctgtttc ttttttcgcc ctatcctact 25980
gctgttctcc ctctgggttac tctttcaaac ctctaattgg tatttttagct ttgatactca 26040
taatctttat taactttata ataattaaca acttacagta aacatgcttg ttctactgct 26100
cgccacatct ttcgctctct ctcacgccag aacaagtatt gttggcgag gttacaatgc 26160
aactcttcaa tctgcttaca tgccagattc cgaccagata ccccatatta cgtgggtactt 26220
acaaacctcc aaacctaatt cttcatttta tgaaggaaac aaactctgcg atgactccga 26280
caacagaacg cacacatttc cccacccttc actacaattc gaatgcgtaa acaaaagctt 26340
gaagctttac aacttaaagc cttcagattc tggcttgtac catgctgtag ttgaaaaaag 26400
taatttagaa gtccacagtg attacattga attgacggtt gtggacctgc cacctccaaa 26460
atgtgagggtt tcctcctctt accttgaagt tcaaggcgtg gatgcctact gcctcataca 26520
cattaactgc agcaactcta aatatccagc tagaatttac tataatggac aggaaagtaa 26580
tcttttttat tatttaacaa caagcgctgg taacggtaaa cagttacctg actattttac 26640
tgctgttggt gaattttcca cctacagaga aacgtatgcc aagcggcctt acaatttctc 26700
atacccgttt aacgacctt gcaatgaaat acaagcgctc gaaactggaa ctgattttac 26760
tccaattttc attgctgcca ttgttgtaag ctttaattacc attattgtca gcctagcatt 26820
ttactgcttt tacaagccca aaaaccctaa gtttgaaaaa cttaaactaa aacctgtcat 26880
tcaacaagtg tgattttggt ttccagcatg gtagctgcat ttctacttct cctctgtcta 26940
cccatcattt tcgtctcttc aactttcgcc gcagtttccc acctggaacc agagtgccta 27000
ccgccttttg acgtgtatct gattctcacc tttgtttggt gtatatccat ttgcagtata 27060
gcctgctttt ttataacaat ctttcaagcc gccgactatt ttacgtgcg aattgcttac 27120
tttagacacc atcctgaata cagaaatcaa aacgttgcct ccttactttg tttggcatga 27180
ttaagttatt gctgatactt aattatttac ccctaataca ctgtaattgt ccattcacca 27240
aaccctggtc attctacacc tggtatgata aaatccccga cactcctgtt gcttggcttt 27300
acgcagccac cgccgctttg gtatttatat ctacttgcct tggagtaaaa ttgtatttta 27360
ttttacacac tgggtggcta catcccagag aagatttacc tagatatcct cttgtaaacy 27420
cttttcaatt acagcctctg cctcctcctg atcttcttcc tcgagctccc tctattgtga 27480

gctactttca actcaccggt ggagatgact gactctcagg acattaatat tagtgtggaa 27540
agaatagctg ctacagcgtca gcgagaaacg cgagtgttgg aatacctgga actacagcaa 27600
cttaaagagt cccactgggtg tgagaaagga gtgctgtgcc atgttaagca ggcagccctt 27660
tcctacgatg tcagcgttca gggacatgaa ctgtcttaca ctttgccttt gcagaaacaa 27720
accttctgca ccatgatggg ctctacctcc atcacaatca cccaacaagc cgggcctgta 27780
gagggggcta tcctctgtca ctgtcacgca cctgattgca tgtccaaact aatcaaaact 27840
ctctgtgctt taggtgatat ttttaagggtg taaatcaata ataaacttac cttaaatttg 27900
acaacaaatt tctgggtgaca tcattcagca gcaccacttt accctcttcc cagctctcgt 27960
atgggatgcg atagtgggtg gcaaacttcc tccaaaccct aaaagaaata ttggtatcca 28020
cttccttgtc ctacccaca attttcatct tttcatagat gaaaagaacc agagtgtgatg 28080
aagacttcaa ccccgctctac ccctatgaca ccacaaccac tcctgcagtt ccctttatat 28140
caccctctt tgtaaacagc gatgggtcttc aggaaaaccc cccaggtgtt ttaagtctgc 28200
gaatagctaa acccctatat ttcgacatgg agagaaaact agccctttca cttggaagag 28260
ggttgacaat taccgccgcc ggacaattag aaagtacgca gagcgtacaa accaaccac 28320
cgttgataat taccaacaac aacacactga ccctacgtca ttctcccccc ttaaacctaa 28380
ctgacaatag cttagtgtca ggctactcga gtcctctccg cgtcacagac aacaaactta 28440
catttaactt cacatcacca ctccgttatg aaaatgaaaa cttactttt aactatacag 28500
agcctcttaa acttataaat aacagccttg ccattgacat caattcctca aaaggcctta 28560
gtagcgtcgg aggctcacta gctgtaaacc tgagttcaga cttaaagttt gacagcaacg 28620
gatccatagc ttttggcata caaacctgt ggaccgctcc gacctcgact ggcaactgca 28680
ccgtctacag cgagggcgat tccctactta gtctctgttt aaccaaattgc ggagctcacg 28740
tcttaggaag tgtaagttta accggtttta caggaaccat aacccaaattg actgatattt 28800
ctgtcaccat tcaatttaca tttgacaaca atggtaagct actaagctct ccacttataa 28860
acaacgcctt tagtattcga cagaatgaca gtacggcctc aaaccctacc tacaacgccc 28920
tggcgtttat gcctaacagt accatatatg caagaggggg aggtggtgaa ccacgaaaca 28980
actactacgt ccaaacgtat cttaggggaa atgttcaaaa accaatcatt cttactgtaa 29040
cctacaactc agtcgccaca ggatattcct tatcttttaa gtggactgct cttgcacgtg 29100
aaaagtttgc aaccccaaca acctcgtttt gctacattac agaacaataa aaccgtgtac 29160
cccaccgttt cgtttttttc agatgaaacg ggcgagagtt gatgaagact tcaaccagt 29220
gtacccttat gacccccac atgctcctgt tatgcccttc attactccac cttttacctc 29280
ctcggatggg ttgcaggaaa aaccacttgg agtggttaagt ttaaaactaca gagatcccat 29340
tactacgcaa aatgagtctc ttacaattaa actaggaaac ggcctcactc tagacaacca 29400
gggacaacta acatcaaccg ctggcgaagt agaacctcca ctactaacg ctaacaacaa 29460
acttgactg gtctatagcg atccttttagc agtaaagcgc aacagcctaa ccttatcgca 29520
caccgctccc cttgttattg ctgataactc ttttagcattg caagtttcag agcctatttt 29580

tataaatgac aaggacaaac tagccctgca aacagccgcg ccccttgtaa ctaacgctgg 29640
cacccttcgc ttacaaagcg ccgccccctt aggcattgca gaccaaacc taaaactcct 29700
gtttaccaac cctttgtact tgcagaataa ctttctcacg ttagccattg aacgaccctt 29760
tgccattacc aatactggaa agctggctct acagctctcc ccaccgctac aaacagcaga 29820
cacaggcttg actttgcaaa ccaacgtgcc attaaactgta agcaacggga ccctaggctt 29880
agccataaag cgccactta ttattcagga caacaacttg tttttggact tcagagctcc 29940
cctgctctt ttcaacagcg acccagtact agggcttaac ttttacacc ctcttgcggt 30000
acgcgatgag gcgctcactg ttaacacagg ccgcggcctc acagtgagtt acgatggttt 30060
aattttaaat cttggttaagg atcttcgctt tgacaacaac accgtttctg tcgctcttag 30120
tgctgctttg cctttacaat aactgatca gcttcgcctt aacgtggcg ctgggctgcg 30180
ttacaatcca gtgagtaaga aattggacgt gaacccaat caaaacaagg gtttaacctg 30240
ggaaaatgac tacctcattg taaagctagg aaatggatta ggttttgatg gcgatggaaa 30300
catagctgtt tctcctcaag ttacatcgcc tgacacctta tggaccactg ccgaccatc 30360
ccccaattgt tccatctaca ctgatttaga tgccaaaatg tggctctcgt tggtaaaaca 30420
agggggtgtg gttcacggtt ctgttgcttt aaaagcattg aaaggaacc tattgagtcc 30480
tacggaaagc gccattgtta ttatactaca ttttgacaat tatggagtgc gaattctcaa 30540
ttatcccact ttgggcactc aaggcacgtt gggaaataat gcaacttggg gttataggca 30600
gggagaatct gcagacacta atgtactcaa tgcactagca tttatgcca gttcaaaaag 30660
gtaccaaga gggcgtggaa gcgaagtcca gaatcaaaact gtgggctaca cttgtataca 30720
gggtgacttt tctatgccg taccgtacca aatacagtac aactatggac caactggcta 30780
ctcctttaaa tttatttggga gaactgtttc aagacaacca tttgacatcc catgctgttt 30840
tttctcttac attacggaag aataaaacaa ctttttcttt ttattttctt tttattttac 30900
acgcacagta aggttcctc cacccttcca tctcacagca tacaccagcc tctccccctt 30960
catggcagta aactgttgtg agtcagtccg gtatttggga gtttaagatcc aaacagtctc 31020
tttggatgat aaacatggat ccgtgatgga cacaaatccc tgggacaggt tctccaacgt 31080
ttcggtaaaa aactgcatgc cgccctacaa aacaaacagg ttcaggctct ccacgggtta 31140
tctccccgat caaactcaga cagagtaaag gtgcgatgat gttccactaa accacgcagg 31200
tggcgctgtc tgaacctctc ggtgcgactc ctgtgaggct ggtaagaagt tagattgtcc 31260
agcagcctca cagcatggat catcagtcta cgagtgcgtc tggcgagca gcgcatctga 31320
atctcactga gattccggca agaatcgac accatcaca ttaggttgtt catgatccca 31380
tagctgaaca cgctccagcc aaagctcatt cgctccaaca gcgccaccgc gtgtccgtcc 31440
aaccttactt taacataaat caggtgtctg ccgctacaa acatgctacc cgcatacaga 31500
acctccggg gcaaaccctt gttcaccacc tgctgtacc agggaaacct cacatttatc 31560
agggagccat agatagccat tttaaacaa ttagctaaca ccgccccacc agctctacac 31620

tgaagagaac cgggagagtt acaatgacag tgaataatcc atctctcata acccctaattg 31680
gtctgatgga aatccagatc taacgtggca cagcagatac acactttcat atacattttc 31740
atcacatgtt tttcccaggc cggtaaaata caatcccaat acacgggcca ctctctgcagt 31800
acaataaagc taataacaaga tggatatactc ctacactcac taacattgtg catgttcata 31860
ttttcacatt ctaagtaccg agagttctcc tctacaacag cactgccgcg gtcctcacia 31920
ggtagtagct ggtgacgatt gtaaggagcc agtctgcagc gataccgtct gtcgcgttgc 31980
atcgtagacc agggaccgac gcacttcctc gtacttgtag tagcagaacc acgtccgctg 32040
ccagcacgtc tccaagtaac gccggtccct gcgtcgctca cgctccctcc tcaacgcaaa 32100
gtgcaaccac tcttgtaatc cacacagatc cctctcggcc tccggggcga tgcacacctc 32160
aaacctacag atgtctcggc acagttccaa acacgtagtg agggcgagtt ccaaccaaga 32220
cagacagcct gatctatccc gacacactgg aggtggagga agacacggaa gaggcattgtt 32280
attccaagcg attcaccaac gggtcgaaat gaagatcccg aagatgacaa cggtcgcctc 32340
cggagccctg atggaattta acagccagat caaacattat gcgattttcc aggctatcaa 32400
tcgcggcctc caaaagagcc tggaccgcga cttccacaaa caccagcaaa gcaaaagcgt 32460
tattatcaaa ctcttcgatc atcaagctgc aggactgtac aatgcccag taattttcat 32520
ttctccactc gcgaatgatg tcgcggcaaa tagtctgaag gttcatgccg tgcataataa 32580
aaagctccga aaggcgcccc tctatagcca tgcgtagaca caccatcatg actgcaagat 32640
atcgggctcc tgagacacct gcagcagatt taacagacct aggtcagggt gctctccgcg 32700
atcgcgaatc tccatccgca aagtcatttg caaataatta aatagatctg cgccgactaa 32760
atctgttaac tccgcgctag gaactaaatc aggtgtggct acgcagcaca aaagttccag 32820
ggatggcgcc aaactcacta gaaccgctcc cgagtagcaa aactgatgaa tgggagtaac 32880
acagtgtaaa atgttcagcc aaaaatcact aagctgctcc tttaaaaagt ccagtacttc 32940
tatattcagt tcgtgcaagt actgaagcaa ctgtgcggga atatgcacag caaaaaaat 33000
agggcggtc agatacatgt tgacctaaaa taaaaagaat cattaaacta aagaagcctg 33060
gcgaacggtg ggatatatga cacgctccag cagcaggcaa gcaaccggct gtccccggga 33120
accgcggtaa aattcatccg aatgattaaa aagaacaaca gagacttccc accatgtact 33180
cggttggatc tcctgagcac agagcaatac cccctcaca ttcatatccg ctacagaaaa 33240
aaaacgtccc agatacccag cgggaatatc caacgacagc tgcaaagaca gcaaaacaat 33300
ccctctggga gcaatcacia aatcctccg tgaaaaaagc acatacatat tagaataacc 33360
ctgttgctgg ggcaaaaagg cccgtcgtcc cagcaaatgc acataaatat gttcatcagc 33420
cattgccccg tcttaccgcg taaacagcca cgaaaaatc gagctaaaat ccaccaaca 33480
gcctatagct atatatacac tccaccaat gacgctaata ccgcaccacc cagacaaaa 33540
gttcaccac acccaciaaa cccgcgaaaa tccagcgccg tcagcacttc cgcaatttca 33600
gtctcaciaa gtcacttccg cgcgcctttt cactttcca cacacgccct tcgcccgcc 33660
gccctcgcgc caccgcgct caccacgt caccgcagc caccgcggcc ccgcctcgt 33720

```

cctccccgct cattatcata ttggcacgtt tccagaataa ggtatattat tgatgcagca 33780
aaacaatccc tctgggagca atcacaaaat cctccggtga aaaaagcaca tacatattag 33840
aataaccctg ttgctggggc aaaaaggccc gtcgtcccag caaatgcaca taaatatgtt 33900
catcagccat tgccccgtct taccgcgtaa acagccacga aaaaatcgag ctaaaatcca 33960
cccaacagcc tatagctata tatacactcc acccaatgac gctaataaccg caccacccac 34020
gaccaaagtt caccacacacc cacaaaaccc gcgaaaatcc agcgccgtca gcacttccgc 34080
aatttcagtc tcacaacgtc acttccgcgc gccttttcac tttccacac acgcccttcg 34140
cccgcccgcc ctgcgcccac ccgcggtcac ccacggtcac cgcacgtcac ccgggccccg 34200
cctcgctcct ccccgctcat tatcatattg gcacgtttcc agaataaggt atattattga 34260
tgca 34264

```

```

<210> 5
<211> 31044
<212> DNA
<213> simian adenovirus SV-25

```

```

<400> 5
catcatcaat aatatacctt attctggaaa cgtgccaata tgataatgag cggggaggag 60
cgaggcgggg ccgggggtgac gtgcggtgac gcgggggtggc gcgagggcgg ggcgaaagggc 120
gcgggtgtgt gtgtgggagg cgcttagttt ttacgtatgc ggaaggagggt tttataccgg 180
aagatgggta atttgggcgt atacttgtaa gttttgtgta atttggcgcg aaaactgggt 240
aatgaggaag ttgagggtta tatgtacttt ttatgactgg gcggaatttc tgctgatcag 300
cagtgaactt tgggcgctga cggggagggt tcgctacgtg acagtaccac gagaaggctc 360
aaaggccccca tttattgtac tcttcagcgt tttcgctggg tatttaaacg ctgtcagatc 420
atcaagaggc cactcttgag tgctggcgag aagagttttc tcctccgtgc tgccacgatg 480
aggctggtcc ccgagatgta cgggtgtttt agcgacgaga cgggtgcgtaa ctcagatgac 540
ctgctgaatt cagacgcgct ggaaatttcc aattcgctg tgctttcgcc gccgtcactt 600
cacgacctgt ttgtgttttg gctcaacgct tagcaacgtg ttatataggg tcaagaagga 660
gcaggagacg cagtttgcta ggctgttggc cgatactcct ggagtttttg tggctctgga 720
tctaggccat cactctcttt tccaagagaa aattatcaaa aacttaactt ttacgtctcc 780
tggtcgcacg gttgcttccg ctgcctttat tacctatatt ttggatcaat ggagcaacag 840
cgacagccac ctgtcgtggg agtacatgct ggattacatg tcgatggcgc tgtggagggc 900
catgctgcgg aggaggggtt gcatttactt gcgggcgcag cctccgcggc tggaccgagt 960
ggaggaggag gacgagccgg gggagaccga gaacctgagg gccgggctgg accctccaac 1020
ggaggactag gtgctgagga tgatcccgaa gaggggacta gtggggctag gaagaagcaa 1080
aagactgagt ctgaacctcg aaactttttg aatgagttga ctgtgagttt gatgaatcgt 1140
cagcgtccgg agacaatttt ctggtctgaa ttggaggagg aattcaggag gggggaactg 1200
aacctgctat acaagtatgg gtttgaacag ttaaaaactc actggttgga gccgtgggag 1260

```

gattttgaaa	ccgccttggga	cacttttgct	aaagtggctc	tgcggccgga	taaggtttac	1320
actatccgcc	gcactgttaa	cataaagaag	agtgtttatg	ttataggcca	tggagctctg	1380
gtgcaggtgc	aaaccgtcga	ccgggtggcc	tttagttgcy	gtatgcaaaa	tctgggcccc	1440
ggggtgatag	gcttaaattg	tgtaacattt	cacaatgtaa	ggtttactgg	tgaaagtttt	1500
aacggctctg	tgtttgcaaa	taacacacag	ctgacgctcc	acggcgttta	cttttttaac	1560
tttaataaca	catgtgtgga	gtcgtggggc	aggggtgtctt	tgaggggctg	ctgttttcac	1620
ggctgctgga	aggcggtggt	gggaagactt	aaaagtgtaa	catctgtaaa	aaaatgcgtg	1680
tttgagcgg	gtgtgttggc	tttaactgtg	gagggtctgtg	gacgcattag	gaataatgcg	1740
gcgtctgaga	atggatgttt	tcttttgcta	aaaggcacgg	ctagtattaa	gcataacatg	1800
atatgcggca	gcggtctgta	cccttcacag	ctgttaactt	gcgcggatgg	aaactgtcag	1860
accttgcgca	ccgtgcacat	agcgtcccac	cagcgccgcg	cctggccaac	attcgagcac	1920
aatatgctta	tgcgttgtgc	cgtccacttg	ggccctaggc	gaggcgtgtt	tgtgccttac	1980
cagtgttaact	ttagccatac	caagatttta	ctagaacctg	ataccttctc	tcgagtgtgt	2040
ttcaatgggg	tgtttgacat	gtcaatggaa	ctgtttaaag	tgataagata	tgatgaatcc	2100
aagtctcggt	gtcgcccatg	tgaatgcgga	gctaatacatc	tgaggttgta	tcctgtaacc	2160
ctaaacgtta	ccgaggagct	gaggacggat	caccacatgt	tgtcctgcct	gcgcaccgac	2220
tatgaatcca	gcgacgagga	gtgaggtgag	gggcggagcc	acaaagggtg	taaaggggcg	2280
tgaggggtgg	gtgtgatgat	tcaaaatgag	cgggacgacg	gacggcaacg	cgtttgaggg	2340
tggagtgttc	agcccttatc	tgacatctcg	tcttccttcc	tgggcaggag	tgcgtcagaa	2400
tgtagtgggc	tccaccgtgg	acggacgacc	ggtcgcccct	gcaaattccg	ccaccctcac	2460
ctatgccacc	gtgggatcat	cggtggacac	tgccgcggca	gctgccgctt	ctgctgccgc	2520
ttctactgct	cgcgcatg	cggtgatatt	tggactgtat	aaccaactgg	ccactgcagc	2580
tgtggcgctt	cggtctctgg	ttcaagaaga	tgccctgaat	gtgatcctga	ctcgccctgga	2640
gatcatgtca	cgctgcttgg	acgaactggc	tgcgcagata	tccaagcta	accccgatac	2700
cacttcagaa	tcctaaaata	aagacaaaca	aatatgttga	aaagtaaaat	ggctttat	2760
gttttttttg	gctcggtagg	ctcgggtcca	cctgtctcgg	tcgttaagaa	ctttgtgtat	2820
gtttttccaa	acacggtaca	gatgggcttg	gatgttcaag	tacatgggca	tgaggccatc	2880
tttgggggtga	agataggacc	attgaagagc	gtcatgctcc	gggggtgggt	tgtaaattac	2940
ccagtcgtag	cagggtttct	gggcgtggaa	ctggaagatg	tccttttagga	gtaggctgat	3000
ggccaagggc	aggcccttag	tgtaggtgtt	tacaaagcgg	ttaagctggg	agggatgcat	3060
gcgggggggag	atgatatgca	tcttggttg	gatcttgagg	ttagctatgt	taccacccag	3120
gtctctgcgg	gggttcatgt	tatgaaggac	caccagcacg	gtgtagccgg	tgcatattggg	3180
gaacttgta	tgcagtttgg	aggggaaggc	gtggaagaat	ttagagaccc	ccttggtggc	3240
ccctaggttt	tccatgcact	catccataat	gatggcaatg	ggacccttg	cggccgcttt	3300

ggcaaacacg ttttgggggt tggaacatc atagttttgc tctagagtga gctcatcata 3360
 ggccatctta acaaagcggg gtaggagggg gcccgactgg gggatgatag ttccatctgg 3420
 gcctggggcg tagttaccct cacagatctg catctcccag gccttaattt ccgagggggg 3480
 tatcatgtcc acctgggggg caataaagaa cacggtttct ggcgggggat tgatgagctg 3540
 ggtggaaagc aagttacgca gcagttgaga tttgccacag ccggtggggc cgtagatgac 3600
 cccgatgacg ggttgacgct ggtagttgag agaggaacag ctgccgtcgg ggcgcaggag 3660
 gggggctacc tcattcatca tgcttctaac atgtttatct tcactcacta agttttgcaa 3720
 gagcctctcc ccacccaggg ataagagttc ttccaggctg ttgaagtgtt tcagcggttt 3780
 taggccgtcg gccatgggca tcttttcgag cgactgacga agcaagtaca gtcggtccca 3840
 gagctcggtg acgtgctcta tggaatctcg atccagcaga cttcttggtt gcgggggttg 3900
 ggtcgacttt cgctgtaggg caccagccgg tgggcgtcca gggccgcgag ggttctgtcc 3960
 ttccagggtc tcagcgtccg ggtgaggggt gtctcgggtga cggtgaaagg atgagccccg 4020
 ggctgggagc ttgcgagggg gcgcttcagg ctcatcctgc tgggtgctgaa gcggacgtcg 4080
 tctccctgtg agtcggccag atagcaacga agcatgaggt cgtagctgag ggactcggcc 4140
 gcgtgtccct tggcgcgcag ctttcccttg gaaacgtgct gacatttggg gcagtgacga 4200
 cattggaggg cgtagagttt gggggccagg aagaccgact cgggcgagta ggcgtcggct 4260
 ccgcactgag cgcagacggt ctgcactcc actagccacg tgagctcggg tttagcggga 4320
 tcaaaaacca agttgcctcc attttttttg atgcgtttct taccttgctt ttccatgagt 4380
 ttgtggcccc cttccgtgac aaaaaggctg tcggtgtctc cgtagacaga cttgaggggg 4440
 cgatcttcca aagggtgtcc gaggtcttcc gcgtacagga actgggacca ctccgagacg 4500
 aaggctctgg tccaggctaa cacgaaggag gcaatctgag aggggtatct gtcgttttca 4560
 atgagggggg ccaccttttc cagggtgtgc agacacaggt cgtcctcctc cggtccacg 4620
 aagggtgatt gcttgtaagt gtaggtcacg tgatctgcac ccccaaagg ggtataaaag 4680
 ggggcgtgcc caccctctcc gtcactttct tccgcatcgc tgtggaccag agccagctgt 4740
 tcgggtgagt aggccctctc aaaagccggc atgatctcgg cgctcaagtt gtcagtttct 4800
 acaaacgagg tggatttgat attcacgtgc cccgcggcga tgcttttgat ggtggagggg 4860
 tccatctgat cagaaaacac gatctttttg ttgtcaagtt tgggtggcgaa agaccgtag 4920
 agggcggttg aaagcaactt ggcgatggag cgcagggtct gatttttctc ccgatcggcc 4980
 ctctccttgg cggcgatgtt gagttgcacg tactcccggg ccgcgcaccg ccactcgggg 5040
 aacacggcgg tgcgctcgtc gggcaggatg cgcacgcgcc agccgcgatt gtgcagggtg 5100
 atgaggtcca cgctggtagc cacctccccg cggaggggct cgttggtcca acacaatcgc 5160
 ccccttttcc tggagcagaa cggaggcagg ggatctagca agttggcggg cggggggctg 5220
 gcgtcgatgg tgaagatacc gggtagcagg atcttattaa aataatcgat ttcggtgtcc 5280
 gtgtcttgca acgcgtcttc ccacttcttc accgccaggg ccctttcgta gggattcagg 5340
 ggcggtcccc agggcatggg gtgggtcagg gccgaggcgt acatgccgca gatgtcatac 5400

acgtacaggg	gttccctcaa	caccccgatg	taagtggggg	aacagcgccc	cccgcggatg	5460
ctggctcgca	cgtagtcgta	catctcgcgc	gagggagcca	tgaggccgtc	tccaagtgg	5520
gtcttgtggg	gtttttcggc	ccggtagagg	atctgtctga	agatggcgtg	ggagtggaa	5580
gagatggtgg	ggcgttgaa	gacgttaaag	ttggccccgg	gtagtccac	ggagtcttgg	5640
atgaactggg	cgtaggattc	ccggagtttg	tccaccaggg	cggcggtcac	cagcacgtcg	5700
agagcgcagt	agtccaacgt	ctcgcggacc	aggttgtagg	ccgtctcttg	tttttctcc	5760
cacagtctgc	ggttgaggag	gtattcctcg	cggctcttcc	agtactcttc	ggcgggaaat	5820
cctttttcgt	ccgctcggtg	agaacctaac	atgtaaaatt	cgttcaccgc	tttgtatgga	5880
caacagcctt	tttctaccgg	cagggcgtag	gcttgagcgg	cctttctgag	agaggtgtgg	5940
gtgagggcga	aggtgtcccc	caccatcact	ttcaggtact	gatgtttgaa	gtccgtgtcg	6000
tcgcaggcgc	cctgttccca	cagcgtgaag	tcggtgcgct	ttttctgcct	gggattgggg	6060
agggcgaagg	tgacatcggt	aaagagtatt	ttcccggcgc	ggggcatgaa	gttgcgagag	6120
atcctgaagg	gcccgggcac	gtccgagcgg	ttgttgatga	cctgcgccgc	caggacgatc	6180
tcgtcgaagc	cgttgatgtt	gtgaccacg	atgtaaagt	cgatgaagcg	cggctgtccc	6240
ttgagggccg	gcgctttttt	caactcctcg	taggtgagac	agtccggcga	ggagagaccc	6300
agctcagccc	gggcccagtc	ggagagttga	ggattagccg	caaggaagga	gctccataga	6360
tccaaggcca	ggagagtttg	caagcggtcg	cggaactcgc	ggaacttttt	ccccacggcc	6420
attttctccg	gtgtcactac	gtaaaagggt	ttggggcggt	tgttccacac	gtcccatcgg	6480
agctctaggg	ccagctcgca	ggcttggcga	acgagggctt	cctcgccaga	gacgtgcatg	6540
accagcataa	agggtagcaa	ctgtttcccc	aacgagccca	tccatgtgta	ggtttctacg	6600
tcgtaggtga	caaagagccg	ctgggtgcgc	gcggtgggagc	cgatcggaag	gaagctgatc	6660
tcctgccacc	agctggagga	atgggtgtta	atgtggtgga	agtagaagtc	ccgccggcgc	6720
acagagcatt	cgtgctgatg	tttgtaaaag	cgaccgcagt	agtcgcagcg	ctgcacgctc	6780
tgtatctcct	gaacgagatg	cgcttttcgc	ccgcgcacca	gaaaccggag	ggggaagttg	6840
agacgggggg	ctggtggggc	gacatcccct	tcgccttggc	ggtgggagtc	tgctgtctcg	6900
tcctccttct	ctgggtggac	gacggtgggg	acgacgacgc	cccgggtgcc	gcaagtccag	6960
atctccgcca	cggaggggtg	caggcgctgc	aggaggggac	gcagctgccc	gctgtccagg	7020
gagtcgaggg	aagtcgcgct	gaggtcggcg	ggaagcgttt	gcaagtccac	tttcagaaga	7080
ccggtgaagag	cgtgagccag	gtgcagatgg	tacttgattt	ccaggggggt	gttgatgaa	7140
gcgtccacgg	cgtagaggag	tccgtgtccg	cgcggggcca	ccaccgtgcc	ccgaggaggt	7200
tttatctcac	tcgtcgaggg	cgagcgccgg	ggggtagagg	cggctctgcg	ccggggggca	7260
gcggaggcag	aggcacgttt	tcgtgaggat	tcggcagcgg	ttgatgacga	gcccggagac	7320
tgctggcgtg	ggcgacgacg	cggcggttga	ggtcctggat	gtgccgtctc	tgctgaaga	7380
ccaccgcccc	ccgggtcctg	aacctaaga	gagttccaca	gaatcaatgt	ctgcatcggt	7440

aacggcgggc	tgcctgagga	tctcctgcac	gtcgcccgag	ttgtcctgat	aggcgatctc	7500
ggccatgaac	tgttccactt	cttcctcgcg	gaggtcaccg	tggcccgctc	gctccacggg	7560
ggcggccagg	tcgttgagga	tgcggcgcat	gagttgagag	aaggcggtga	ggccggtctc	7620
gttccacacg	cggctgtaca	ccacgtttcc	gaaggagtcg	cgcgctcgca	tgaccacctg	7680
ggccacgttg	agttccacgt	ggcgggcgaa	gacggcgtag	tttctgaggc	gctggaagag	7740
gtagttgagc	gtggtggcga	tgtgctcgca	gacgaagaag	tacataatcc	agcgccgcag	7800
ggatcatctc	ttgatgtctc	cgatggcttc	gagacgctcc	atggcctcgt	agaagtcgac	7860
ggcgaagtgt	aaaaattggg	agttgcgggc	ggccaccgtg	agttcttctt	gcaggaggcg	7920
gatgagatcg	gcgaccgtgt	cgcgcacctc	ctgttcgaaa	gcgccccgag	gcgcctctgc	7980
ttcttctctc	ggctcctcct	cttcaggggg	ctcgggttcc	tccggcagct	ctgcgacggg	8040
gacggggcgg	cgacgtcgtc	gtctgaccgg	caggcggtcc	acgaagcgct	cgatcatttc	8100
gccgcgccgg	cgacgcatgg	tctcggtgac	ggcgcgtccg	ttttcgcgag	gtcgcagttc	8160
gaagacgccg	ccgcgcagag	cgcggcggtg	caggaggagg	aagtgggttag	ggccgctcggg	8220
cagggaacacg	gcgctgacga	tgcattttat	caattgctgc	gtaggcactc	cgtgcaggga	8280
tctgagaacg	tcgaggtcga	cgggatccga	gaacttctct	aggaaagcgt	ctatccaatc	8340
gcaatcgcaa	ggtaagctga	gaacgggtggg	tcgctggggg	gcgttcgcgg	gcagttggga	8400
ggtgatgctg	ctgatgatgt	aattaaagta	ggcggctctc	aggcggcgga	tggtggcgag	8460
gaggaccacg	tctttggggc	cggcctgttg	aatgcgcagg	cgctcggcca	tgccccaggc	8520
ctcgtctctg	cagcgacgca	ggtctttgta	gaagtcttgc	atcagtctct	ccaccggaac	8580
ctctgcttct	cccctgtctg	ccatgcgagt	cgagccgaac	ccccgcaggg	gctgcagcaa	8640
cgctaggtcg	gccacgaccc	tttcggccag	cacggcctgt	tgaatctgcg	tgagggtggc	8700
ctggaagtcg	tccaggtcca	cgaagcgggt	ataggccccc	gtgttgatgg	tgtaggtgca	8760
gttgggccatg	acggaccagt	tgacgacttg	catgccgggt	tgggtgatct	ccgtgtactt	8820
gaggcgcgag	taggccctgg	actcgaacac	gtagtcgttg	catgtgcgca	ccagatactg	8880
gtagccgacc	aggaagttag	gaggcggtc	tcggtacagg	ggccagccaa	cgggtggcggg	8940
ggcgccgggg	gacaggtcgt	ccagcatgag	gcggtggtag	tggtagatgt	agcgggagag	9000
ccaggtgatg	ccggccgagg	tggttgcggc	cctggtgaat	tcgcggacgc	ggttccagat	9060
gttgcgcgagg	ggaccaaagc	gctccatggt	gggcacgctc	tgccccgtga	ggcggggcgca	9120
atcttgtagc	ctctagatgg	aaaaaagaca	gggcgggtcat	cgactccttt	ccgtagcttg	9180
gggggttaaag	tcgcaagggt	gcggcgggcg	ggaaccccg	ttcgagaccg	gccggatccg	9240
ccgctcccga	tgcgctggc	cccgcaccca	cgacgtccgc	gccgagaccc	agccgcgacg	9300
ctccgccccca	atacggaggg	gagtcctttg	gtgttttttc	gtagatgcat	ccggtgctgc	9360
ggcagatgag	acccagacg	cccactacca	ccgccgtggc	ggcagtaaac	ctgagcggag	9420
gcgggtgacag	ggaggaggaa	gagctggctt	tagacctgga	agagggagag	gggctggccc	9480
ggctgggagc	gcatcccca	gagagacacc	ctagggttca	gctcgtgagg	gacgccaggc	9540

aggcttttgt	gccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	9600
tgcgcgattg	caggtttcgg	gcgggcagag	agctcagggc	gggcttcgat	cgggagcggc	9660
tcctgagggc	ggaggatttc	gagcccgcag	agcgttcttg	ggtgagcccc	gcccgcgctc	9720
acgtatcggc	ggccaacctg	gtgagcgcgt	acgagcagac	ggtgaacgag	gagcgcaact	9780
tccaaaagag	ctttaacaat	cacgtgagga	ccctgatcgc	gagggaggag	gtgaccatcg	9840
ggctgatgca	tctgtgggac	ttcgtggagg	cctacgtgca	gaacccggct	agcaaacccc	9900
tgacggccca	gctgttcctg	atcgtgcagc	acagccgcga	caacgagacg	ttccgcgacg	9960
ccatgttgaa	catcgcggag	cccgaggggtc	gctggctctt	ggatctgatt	aacatcctgc	10020
agagcatcgt	ggtgcaggag	aggggcctga	gtttagcgga	caaggtggcg	gccattaact	10080
attcgatgca	gagcctgggg	aagttctacg	ctcgcaagat	ctacaagagc	ccttacgtgc	10140
ccatagacaa	ggagggtgaag	atagacagct	tttacatgcg	catggcgctg	aaggtgctga	10200
cgctgagcga	cgacctcggc	gtgtaccgta	acgacaagat	ccacaaggcg	gtgagcgcca	10260
gccgccggcg	ggagctgagc	gacagggagc	tgatgcacag	cctgcagagg	gcgctggcgg	10320
gcgccgggga	cgaggagcgc	gaggcttact	tcgacatggg	agccgatctg	cagtggcgctc	10380
ccagcgcgcg	cgcccttgag	gcggcggggt	atcccgcgca	ggaggatcgg	gacgatttgg	10440
aggaggcagg	cgagtacgag	gacgaagcct	gaccgggcag	gtgttggttt	agatgcagcg	10500
gccggcggac	gggaccaccg	cggatcccg	acttttgcca	tccatgcaga	gtcaaccttc	10560
gggcgtgacc	gcctccgatg	actgggcggc	ggccatggac	cgcatcatgg	cgctgaccac	10620
ccgcaacccc	gaggctttta	ggcagcaacc	ccaggccaac	cgtttttcgg	ccatcttgga	10680
agcggtggtg	ccgtcgcgca	ccaacccgac	gcacgagaaa	gtcctgacta	tcgtgaacgc	10740
cctggtagac	agcaaggcca	tccgccgtga	cgaggcgggc	ttgatttaca	acgctctttt	10800
ggaacgcgtg	gcgcgctaca	acagcactaa	cgtgcagacc	aatctggacc	gcctcaccac	10860
cgacgtgaag	gaggcgctgg	cgcagaagga	gcggtttctg	agggacagta	atctgggctc	10920
tctggtggca	ctgaacgcct	tcctgagctc	acagccggcc	aacgtgcccc	gcgggcagga	10980
ggattacgtg	agcttcatca	gcgctctgag	actgctggtg	tccgaggtgc	cccagagcga	11040
ggtgtaccag	tctgggccgg	attacttttt	ccagacgtcc	cgacagggct	tgcaaacggt	11100
gaacctgact	caggccttta	aaaacttgca	aggcatgtgg	ggggtcaagg	ccccggtggg	11160
cgatcgcgcc	actatctcca	gtctgctgac	ccccaacact	cgcctgctgc	tgctcttgat	11220
cgcaccgttt	accaacagta	gcactatcag	ccgtgactcg	tacctgggtc	atctcatcac	11280
tctgtaccgc	gaggccatcg	gccaggctca	gatcgacgag	catacgtatc	aggagattac	11340
taacgtgagc	cgtgccctgg	gtcaggaaga	taccggcagc	ctggaagcca	cgttgaactt	11400
tttgctaacc	aaccggaggc	aaaaaatacc	ctcccagttc	acgttaagcg	ccgaggagga	11460
gaggattctg	cgatacgtgc	agcagtccgt	gagcctgtac	ttgatgcgcg	agggcgccac	11520
cgcttccacg	gctttagaca	tgacggctcg	gaacatggaa	ccgtcctttt	actccgcca	11580

ccggccgttc attaacccgtc tgatggacta cttccatcgc gcggccgccca tgaacgggga 11640
 gtacttcacc aatgccatcc tgaatccgca ttggatgccc ccgtccggct tctacaccgg 11700
 ggagtttgac ctgcccgaag ccgacgacgg ctttctgtgg gacgacgtgt ccgatagcat 11760
 ttccacgccg gctaatacgcc gattccagaa gaaggagggc ggagacgagc tccccctctc 11820
 cagcgtggaa gcggcctcaa ggggagagag tccctttcca agtctgtctt ccgccagtag 11880
 cggtcgggta acgcgtccac ggttgccggg ggagagcgac tacctgaacg accccttgct 11940
 gcgaccggct agaaagaaaa attttcccaa taacggggtg gaaagcttgg tggataaaat 12000
 gaatcgttgg aagacgtacg cccaggagca gcgggagtgg gaggacagtc agccgcggcc 12060
 gctggtaccg ccgcattggc gtcgccagag agaagaccg gacgactccg cagacgatag 12120
 tagcgtgttg gacctgggag ggagcggagc caacccttt gctcacttgc aaccaaggg 12180
 gcgctcgagt cgctgtatt aataaaaaag acgcggaaac ttaccagagc catggccaca 12240
 gcgtgtgtgc tttcttctc tctttcttcc tcggcgcggc agaatagaaa gagcgggtgag 12300
 agtcacgccg gcggcgtatg agggcccgc cccttcttac gaaagcgtga tgggatcagc 12360
 gaacgtgccg gccacgctgg aggcgcctta cgttctctcc agatacctgg gacctacgga 12420
 gggcagaaac agcatccgtt actccgagct ggcgcccctg tacgatacca ccaaggtgta 12480
 cctggtggac aacaagtcgg cggacatcgc ctccctgaat taccaaaacg atcacagtaa 12540
 ctttctgact accgtggtgc agaacaatga cttcaccccg acggaggcgg gcacgcagac 12600
 cattaacttt gacgagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcaacacaa 12660
 catgcccac atcaacgagt tcatgtccac caacaagttc agggctaagc tgatggtaga 12720
 aaaaagtaat gcggaaactc ggcagccccg atacgagtgg ttcgagttta ccattccaga 12780
 gggcaactat tccgaaacta tgactatcga tctcatgaat aacgcgatcg tggacaatta 12840
 cctgcaagtg gggagacaga acggggtgct ggaaagcgat atcggcgtga aattcgatac 12900
 cagaaacttc cgactggggt gggatcccg gaccaagctg gtgatgccag gcgtgtacac 12960
 caacgaggct tttcacccgg acatcgtgct gctgccgggg tgcggtgtgg acttcactca 13020
 gagccgtttg agtaacctgt taggaattag aaagcgccgc cccttccaag agggctttca 13080
 aatcatgtat gaggacctgg agggaggtaa tatacccgcc ttactggacg tgtcgaagta 13140
 cgaagctagc atacaacgcg ccaaagcgga gggtagagag attcggggag acacctttgc 13200
 ggtagctccc caggacctgg aaatagtgcc tttaactaaa gacagcaaag acagaagcta 13260
 caatattata aacaacacga cggacaccct gtatcgagc tggtttctgg cttacaacta 13320
 cggagacccc gagaaaggag tgagatcatg gaccatactc accaccacgg acgtgacctg 13380
 tggctcgcag caagtgtact ggtccctgcc ggatatgatg caagaccgg tcaccttcg 13440
 cccctccacc caagtcagca acttcccggg ggtgggcacc gagctgctgc ccgtccatgc 13500
 caagagcttc tacaacgagc aggccgtcta ctcgcaactt attcgccagt ccaccgcgct 13560
 taccacgtg ttcaatcgct ttcccagaa ccagattctg gtgcgccctc ccgtcctac 13620
 cattaccacc gtcagtgaac acgttcccgc cctcacagat cacggaacct tgccgctgcg 13680

cagcagtatc agtggagttc agcgcgtgac catcaccgac gccagacgtc gaacctgccc 13740
ctacgtttac aaagcgcttg gcgtggtggc tcctaaagtt ctttctagtc gcaccttcta 13800
aaaacatgtc catcctcatc tctcccgata acaacaccgg ctggggactg ggctccggca 13860
agatgtacgg cggagccaaa aggcgctcca gtcagcacc agttcgagtt cggggccact 13920
tccgcgctcc ttggggagct tacaagcgag gactctcggg tcgaacggct gtagacgata 13980
ccatagatgc cgtgattgcc gacgcccgcc ggtacaaccc cggaccggtc gctagcgccg 14040
cctccaccgt ggattccgtg atcgacagcg tggtagccgg cgctcgggcc tatgctcgcc 14100
gcaagaggcg gctgcatcgg agacgtcgcc ccaccgccgc catgctggca gccagggccg 14160
tgctgaggcg ggcccggagg gcaggcagaa gggctatgcy ccgcgctgcc gccaacgccg 14220
ccgccgggag ggcccgccga caggctgccc gccaggctgc cgctgccatc gctagcatgg 14280
ccagaccag gagagggaaac gtgtactggg tgctgattc tgtgacggga gtccgagtgc 14340
cgggtgcgag ccgacctccc cgaagttaga agatccaagc tgcgaagacg gcggtactga 14400
gtctccctgt tgttatcagc ccaacatgag caagcgcaag tttaaagaag aactgctgca 14460
gacgctggtg cctgagatct atggccctcc ggacgtgaag ccagacatta agccccgcga 14520
tatcaagcgt gttaaaaagc gggaaaagaa agaggaactc gcggtggtag acgatggcgg 14580
agtggaaatt attaggagtt tcgccccgcg acgcagggtt caatggaaag ggcggcgggt 14640
acaacgcgtt ttgaggccgg gcaccgcggt agtttttacc ccgggagagc ggtcggccgt 14700
taggggtttc aaaaggcagt acgacgaggt gtacggcgac gaggacatat tggaacaggc 14760
ggctcaacag atcggagaat ttgcctacgg aaagcgttcg cgctcggaag acctggccat 14820
cgccttagac agcggcaacc ccacgccag cctcaaacc gtgacgctgc agcagggtgct 14880
tcccgtgagc gccagcacgg acagcaagag ggggattaag agagaaatgg aagatctgca 14940
tcccaccatc caactcatgg tccctaaacg gcagaggctg gaagaggctc tggagaagat 15000
gaaagtggac ccagcatag agccggatgt aaaagtcaga cctattaagg aagtggcccc 15060
cggctcttggg gtgcaaacgg tggacattca aatccccgtc accaccgctt caaccgccgt 15120
ggaagctatg gaaacgcaaa cggagacccc tgccgcgac ggtaccaggg aagtggcgtt 15180
gcaaacggag ccttggtacg aatacgcagc ccctcggcgt cagaggcggt cgcctcgta 15240
cggccccgcc aacgccatca tgccagaata tgcgctgcat ccgtctattc tgccactcc 15300
cggataccgg ggtgtgacgt atcgcccgtc tggaacccgc cgccgaaccc gtcgccgccg 15360
ccgctcccgt cgcgctctgg ccccgtgtc ggtgcggcgt gtgaccgcc ggggaaagac 15420
agtcgtcatt cccaaccgcg gttaccacc tagcatcctt taataactct gccgttttgc 15480
agatggctct gacttgccgc gtgcgccttc ccgttcgcga ctatcgagga agatctcgtc 15540
gtaggagagg catgacgggc agtggtcgcc ggcgggcttt gcgcaggcg atgaaaggcg 15600
gaattttacc cgccctgata ccataattg ccgccgccat cggtgccata cccggcggtt 15660
cttcagtggc gttgcaagca gctcgtaata aataaacaaa ggcttttgca cttatgacct 15720

ggtcctgact attttatgca gaaagagcat ggaagacatc aattttacgt cgctggctcc 15780
 gcggcacggc tcgcggccgc tcatgggcac ctggaacgac atcggcacca gtcagctcaa 15840
 cgggggcgct ttcaattggg ggagcctttg gagcggcatt aaaaactttg gctccacgat 15900
 taaatcctac ggagcaaag cctggaacag tagtgctggg cagatgctcc gagataaact 15960
 gaaggacacc aacttccaag aaaaagtggg caatgggggtg gtgaccggca tccacggcgc 16020
 ggtagatctc gccaaccaag cgggtgcagaa agagattgac aggcgttttg aaagctcgcg 16080
 ggtgccgccc cagagagggg atgaggtgga ggtcgaggaa gtagaagtag aggaaaagct 16140
 gccccgctg gagaaagttc ccggtgcgcc tccgagaccg cagaagcggc ccaggccaga 16200
 actagaagag actctggtga cggagagcaa ggagcctccc tcgtacgagc aagccttgaa 16260
 agagggcgcc tctccaccct cctacccgat gactaagccg atcgcaccca tggctcgacc 16320
 ggtgtacggc aaggattaca agcccgtcac gctagagctg cccccaccgc cccccacgcg 16380
 cccgaccgtc cccccctgc cgactccgtc ggcgccgcg gcgggaccgc tgtccgcacc 16440
 atccgctgtg cctctgccag ccgcccgtcc agtggccgtg gccactgcca gaaaccccag 16500
 aggccagaga ggagccaact ggcaaagcac gctgaacagc atcgtgggccc tgggagtga 16560
 aagcctgaaa cgccgccgtt gctattatta aaaaagtgtg gctaaaaagt ctcccgttgt 16620
 atacgcctcc tatgttaccg ccagagacga gtgactgtcg ccgcgagcgc cgctttcaag 16680
 atggccaccc catcgatgat gccgcagtgg tcttacatgc acatcgccgg ccaggacgcc 16740
 tcggagtacc tgagtcccgg cctcgtgcag tttgcccgcg ccaccgacac ctacttcagc 16800
 ttgggaaaca agtttagaaa ccccaccgtg gccccaccc acgatgtgac cacggaccgc 16860
 tcgcagaggc tgaccctgcg ctttgtgccc gtagaccggg aggacaccgc gtactcttac 16920
 aaagtgcgt acacgttggc cgtaggggac aaccgagtgc tggacatggc cagcacctac 16980
 tttgacatcc ggggggtgct ggatcggggt cccagcttca agccctattc cggcaccgct 17040
 tacaactccc tggcccccaa gggagctccc aaccctcgg aatggacgga cacttccgac 17100
 aacaaactta aagcatatgc tcaggctccc taccagagtc aaggacttac aaaggatggg 17160
 attcaggttg ggctagttgt gacagagtca ggacaaacac cccaatatgc aaacaaagt 17220
 taccaacccg agccacaaat tggggaaaac caatggaatt tagaacaaga agataaagcg 17280
 gcgggaagag tcctaaagaa agatacccct atgtttccct gctatgggtc atatgccagg 17340
 cccacaaacg aacaaggagg gcaggcaaaa aaccaagaag tagatttaca gttttttgcc 17400
 actccgggcg acaccagaa cacggctaaa gtggtacttt atgctgaaaa tgtcaacctg 17460
 gaaactccag atactcactt agtgtttaaa cccgatgacg acagcaccag ttcaaaactt 17520
 cttcttgggc agcaggctgc acctaacaga cccaactaca taggttttag agataatttt 17580
 attggtttta tgtactacaa tagcactgga aacatgggcg tgctggccgg acaggcttct 17640
 caattgaatg ccgtagtcga cttgcaggac agaaacaccg agttgtccta ccagctgatg 17700
 ctggacgcac tgggggatcg cagccgatat ttttcaatgt ggaatcaggc agtagacagc 17760
 tatgaccag acgttagaat tatagaaaac cacggagtgg aagacgaact gccaaactat 17820

tgttttcctc tgggaggaat ggtggtgact gacaattaca actctgtgac gcctcaaaat 17880
 ggaggcagtg gaaatacatg gcaggcagac aatactacat ttagtcaaag aggagcgcag 17940
 attggctccg gaaacatggt tgccctggaa attaacctac aggccaacct ctggcgcggc 18000
 ttcttgattt ccaatattgg gttgtatctt ccagactctc tgaaaatcac ccccgacaac 18060
 atcacgctgc cagaaaacaa aaacacttat cagtacatga acggtcgcgt aacgccaccc 18120
 gggctcatag acacctatgt aaacgtgggc gcgcgctggt ccccgatgt catggacagc 18180
 attaacccct tcaaccacca ccgtaacgcg ggcttgcgct accgctccat gctcttgggc 18240
 aacggccggt atgtgccttt tcacattcag gtgccccaaa aattctttgc cattaanaac 18300
 ctgctgcttc tccccggttc ctatacctat gagtggaact tccgcaagga tgtcaacatg 18360
 atcctgcaga gctcgtgagg taatgacctg cgagtggacg gggccagcat acgctttgac 18420
 agcattaacc tgtatgcca cttttttccc atggcccaca acacggcctc taccctggaa 18480
 gccatgctgc gcaacgacac caatgaccag tccttcaacg actacctgtg cgcggttaac 18540
 atgctgtacc ccatccccgc caacgccacc agcgtgcca tttctattcc ttctcggaac 18600
 tgggctgcct tcaggggctg gagttttact cgctcaaaa ccaaggagac tccctcgtg 18660
 ggctccggtt ttgacccta ctttgtttac tccggctcca ttccctacct agatggcacc 18720
 ttttacctca accacacttt caaaaagggtg tctattatgt ttgactcctc ggtagctgg 18780
 cccggcaacg accgctgct aacgccaac gagttcgaaa ttaagcgttc cgtggacggt 18840
 gaagggtaca acgtggcca gagcaacatg accaaggact ggtttctaata tcaaatgctc 18900
 agtcactata atataggtta ccagggttc tatgtgcccg agaactacaa ggaccgcatg 18960
 tactccttct tccgcaactt ccaaccaatg agccggcagg tggtagatac cgtgacttat 19020
 acagactaca aagatgtcaa gctcccctac caacacaaca actcagggtt cgtgggctac 19080
 atgggaccca ccatgcgaga gggacaggcc taccggcca actatcccta cccctgatc 19140
 ggagagactg ccgtaccag cctcacgcag aaaaagtcc tctgcgaccg ggtgatgtgg 19200
 aggataccct tctctagcaa ctttatgtcg atgggtccc tcaccgacct ggggcagaac 19260
 atgctgtacg ccaactccgc tcacgccttg gacatgactt ttgagggtga tcccatggat 19320
 gagccacgc ttctctatgt tctgtttgaa gtcttcgacg tgggtgcgcat ccaccagccg 19380
 caccgcgcg tcatcgaggc cgtctacctg cgcacacctt tctctgccg taacgccacc 19440
 acctaaagaa gctgatgggt tccagcgaac aggagttgca ggccattgtt cgcgacctgg 19500
 gctgcggggc ctgctttttg ggcaccttcg acaagcgttt tcccgattc atgtccccc 19560
 acaagccggc ctgcgccatc gttaacacgg ccggacggga gacagggggg gtgcactggc 19620
 tcgccttcgc ctggaacccg cgcaaccgca cctgctacct gttcgaccct tttggtttct 19680
 ccgacgaaag gctgaagcag atctaccaat tcgagtacga ggggtcctc aagcgacgcg 19740
 ctctggcctc cagcccgac cactgctca ccctggaaaa gtccaccag acggtccagg 19800
 ggcccctctc ggccgcctgc gggcttttct gttgcatgtt ttgacgccc ttcgtgcact 19860

ggccctcacac ccccatggag cgcaacccca ccatggatct gctcaccgga gtgcccaca 19920
 gcatgcttca cagtccccag gtcgccccca ccttgcgtcg caatcaggac cacctgtatc 19980
 gctttctggg gaaacactct gcctatttcc gccgccaccg gcagcgcacg gaacaggcca 20040
 cggccttcga aagcatgagc caaagagtgt aatcaataaa aaccgttttt atttgacatg 20100
 atacgcgctt ctggcgTTTT tattaataaat cgaagggttc gagggagggg tcctcgtgcc 20160
 cgctggggag ggacacgttg cggacttgga atcgggcgct ccaacgaaac tcggggatca 20220
 ccagccgagg cagggccacg tcttccatgt tctgcttcca aaactgtcgc accagctgca 20280
 gggctcccat cacgtcgggc gctgagatct tgaagtcgca gttagggccg gagccccgcg 20340
 ggctgttgcg gaacacgggg ttggcacact ggaacaccaa cacgtcgggg ttgtggatac 20400
 tagccagggc cgtcgggtcg gtcacctcg atgcatccag atcctcggca ttgctcaggg 20460
 cgaacggggg cagcttgac atctgccgcc cgatctgggg taccaggctc cgcttggtga 20520
 ggagtcgca gcgcagagg atgaggatgc gacgtgccc gcgttgcatg atggggtaac 20580
 tcgccccag gaactcctct atctgacgga aggccatctg ggccttgacg ccctcgggtga 20640
 aaaatagccc acaggacttg ctggaaaaca cgttattgcc acagttgatg tcttccgcgc 20700
 agcagcgcgc atcttcgttc ttgagctgaa ccacgttgcg accccagcgg ttctgaacca 20760
 ccttggtttt cgtggggtgc tcttcagcg cccgctgtcc gttctcgtg gtcacatcca 20820
 tttccaccac gtgtccttg cagaccatct ccactccgtg gaaacagaac agaatgccct 20880
 cctgttgggt attgcgatgc tcccacacgg cgacccggg ggactcccag ctcttggtt 20940
 tcacccccgc gtaggcttcc atgtaagcca ttagaaatct gcccatcagc tcagtgaagg 21000
 tcttctggtt ggtgaagggt agcggcaggc cgcggtgttc ctcgttcaac caagtttgac 21060
 agatcttgcg gtacacggct ccctggtcgg gcagaaactt aaaagtcgtt ctgctctcgt 21120
 tgtccacgtg gaacttctcc atcaacatcg tcatgacttc catgcccttc tcccaggcag 21180
 tcaccagcgg cgcgctctcg gggttcttca ccaacacggc ggtggagggg ccctcgccgg 21240
 ccccgacgtc cttcatggac attttttgaa actccacggg gccgtccgcg cggcgctactc 21300
 tgcgcatcgg agggtagctg aagccacct ccatgacggg gctttcgccc tcgctgtcgg 21360
 agacgatctc cggggagggc ggcggaacgg gggcagactt gcgagccttc ttcttgggag 21420
 ggagcggagg cacctcctgc tcgcgctcgg gactcatctc ccgcaagtag ggggtgatgg 21480
 agcttccttg ttggttctga cggttggcca ttgtatccta ggcagaaaga catggagctt 21540
 atgcgcgagg aaactttaac cgccccgtcc cccgtcagcg acgaagagg catcgtcgaa 21600
 caggacccgg gctacgttac gccgccgag gatctggagg ggcccttaga cgaccggcgc 21660
 gacgctagtg agcggcagga aaatgagaaa gaggaggagg agggctgcta cctcctggaa 21720
 ggcgacgttt tgctaaagca tttcgccagg cagagcacca tactcaagga ggccttgcaa 21780
 gaccgctccg aggtgccctt ggacgtcgcc gcgctctccc aggcctacga ggcgaacctt 21840
 ttctcgcccc gagtgccctc gaagagacag cccaacggca cctgcgagcc caaccgcga 21900
 ctcaacttct acccgtgtt cgccgtgccc gaggcgctgg ccacctacca catctttttc 21960

aaaaaccagc	gcattccccct	ttcctgccgg	gccaaaccgca	ccgcggccga	taggaagcta	22020
acactcagaa	acggagtcag	catacctgat	atcacgtcac	tggaggaagt	gcctaagatc	22080
ttcgaggggtc	tgggtcgaga	tgagaagcgg	gcggcgaacg	ctctgcagaa	agaacagaaa	22140
gagagtcaga	acgtgctggt	ggagctggag	ggggacaacg	cgcgcttgac	cgtcctcaaa	22200
cgttgcatag	aagtttccca	cttcgcctac	ccggccctca	acctgccgcc	caaagttatg	22260
aaatcgggtca	tggaccagct	actcatcaag	agagctgagc	ccctgaatcc	cgaccaccct	22320
gaggcggaaa	actcagagga	cggaaagccc	gtcgtcagcg	acgaggagct	cgagcgggtgg	22380
ctggaaacca	gggacccccca	gcagttgcaa	gagaggcgca	agatgatgat	ggcggccgtg	22440
ctgggtcacgg	tggagctaga	atgcctgcaa	cggtttttca	gcgacgtgga	gacgctacgc	22500
aaaatcgggg	agtccctgca	ctacaccttc	cgccagggct	acgttcgccca	ggcctgcaaa	22560
atctccaacg	tagagctcag	caacctgggt	tcctacatgg	gcacccctca	cgagaaccgg	22620
ctggggcaga	gcgtgctgca	ctgcaccttg	caaggcgagg	cgcgaaagga	ctacgtccga	22680
gactgcgtct	acctcttcct	caccttcacc	tggcagaccg	ccatgggcgt	gtggcagcag	22740
tgcttggaag	agagaaacct	caaagagctg	gacaaactcc	tctgccgcca	gcggcggggcc	22800
ctctggaccg	gcttcagcga	gcgcacgggtc	gcctgcgccc	tggcagacat	cattttccca	22860
gaacgcctga	tgaaaacctt	gcagaacggc	ctgccggatt	tcacagtcga	gagcatcttg	22920
caaaacttcc	gctccttcgt	cctggagcgc	tccgggatct	tgcccgccat	gagctgcgcg	22980
ctgccttctg	actttgtccc	cctttcctac	cgcgagtgcc	ctccccact	gtggagccac	23040
tgctacctct	tccaactggc	caactttctg	gcctaccact	ccgacctcat	ggaagacgtg	23100
agcggagagg	ggctgctcga	gtgccactgc	cgctgcaacc	tctgcacccc	ccacagatcg	23160
ctggcctgca	acaccgagct	gctcagcgaa	acccaggtca	taggtacctt	cgagatccag	23220
gggccccagc	agcaagaggg	tgcttccggc	ttgaagctca	ctccggcgct	gtggacctcg	23280
gcttacttac	gcaaatttgt	agccgaggac	taccacgccc	acaaaattca	gttttacgaa	23340
gaccaatctc	gaccaccgaa	agccccctc	acggcctgcg	tcacaccca	gagcaaaatc	23400
ctggcccaat	tgcaatccat	caaccaagcg	cgccgagatt	tccttttgaa	aaagggtcgg	23460
ggggtgtacc	tggacccccca	gaccggcgag	gaactcaacc	cgtccacact	ttccgtcgaa	23520
gcagcccccc	cgagacatgc	cacccaaggg	aaccgccaag	cagctgatcg	ctcggcagag	23580
agcgaagaag	caagagctgc	tccagcagca	ggtggaggac	gaggaagagc	tgtgggacag	23640
ccaggcagag	gaggtgtcag	aggacgagga	ggagatggaa	agctgggaca	gcctagacga	23700
ggaggacgag	ctttcagagg	aagaggcgac	cgaagaaaaa	ccacctgcat	ccagcgcgcc	23760
ttctctgagc	cgacagccga	agccccggcc	cccagcggcc	ccggccggct	cactcaaagc	23820
cagccgtagg	tgggacgcca	ccggatctcc	agcggcagcg	gcaacggcag	cgggtaaggc	23880
caaacgcgag	cggcgggggt	attgctcctg	gcggaccac	aaaagcagta	tcgtgaactg	23940
cttgcaacac	tgcgggggaa	acatctcctt	tgcccagcgc	tacctcctct	tccatcacgg	24000

tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24060
 gctcggagaa aaaagctaag gcctcctctg ccgcgaggaa aaactccgcc gccgctgccg 24120
 ccaaggatcc gccggccacc gaggagctga gaaagcgcat ctttcccact ctgtatgcta 24180
 tctttcagca aagccgcggg cagcacccctc agcgcgaact gaaaataaaa aaccgctcct 24240
 tccgctcact caccgcagc tgtctgtacc acaagagaga agaccagctg cagcgcaccc 24300
 tggacgacgc cgaagcactg ttcagcaaat actgctcagc gtctcttaaa gactaaaaga 24360
 cccgcgcttt ttccccctcg ggcgccaaaa ccacgctcat cgccagcatg agcaaggaga 24420
 ttcccacccc ttacatgtgg agctatcagc ccagatggg cctggccgcg ggggccgccc 24480
 aggactactc cagcaaaatg aactggctca gcgccggccc ccacatgatc tcacgagtta 24540
 acggcatccg agcccaccga aaccagatcc tcttagaaca ggcggaatc accgccacac 24600
 cccggcgcca actcaacccg ccagttggc ccgccgcca ggtgtatcag gaaactcccc 24660
 gcccgaccac agtcctcctg ccacgcgacg cggaggccga agtcctcatg actaactctg 24720
 gggtagaatt agcgggcccggg tccaggtacg ccaggtacag aggtcgggccc gctccttact 24780
 ctcccgggag tataaagagg gtgatcattc gaggccgagg tatccagctc aacgacgagg 24840
 cggtagctc ctcaaccggt ctcagacctg acggagtctt ccagctcggg ggagcgggccc 24900
 gctcttcctt caccactcgc caggcctacc tgaccctgca gagctcttcc tcgcagccgc 24960
 gctccggggg aatcggcact ctccagttcg tggaagagtt cgtcccctcc gtctacttca 25020
 acccgtttcc cggctcacct ggacgctacc cggacgcctt cattccaac ttgacgcag 25080
 tgagtgaatc cgtggacggc tacgactgat gacagatggg gcggccgtga gagctcggct 25140
 gcgacatctg catcactgcc gccagcctcg ctgctacgct cgggaggcga tcgtgttcag 25200
 ctactttgag ctgccggacg agcaccctca gggaccggct cacgggttga aactcgagat 25260
 tgagaacgcg cttgagtctc acctcatcga cgccttcacc gcccggcctc tcctggtaga 25320
 aaccgaacgc gggatcacta ccatcacctt gttctgcatc tgccccacgc ccggattaca 25380
 tgaagatctg tgttgtcatc ttgcgctca gtttaataaa aactgaactt ttgcccgtac 25440
 cttcaacgcc acgcgttggt tctccttggt aaaaaacccc aggagtctt aacttacaca 25500
 tagcaaaacc cttgtatttt accatagaaa aacaactagc ctttcaatt ggaaaagggt 25560
 taacaatttc tgctacagga cagttggaaa gcacagcaag cgtacaggac agcgctacac 25620
 caccctacg tggatattcc cttttaaagc tgacagacaa cggtttaaca ttaagctatt 25680
 cagatcccct gcgtgtggta ggtgaccaac ttacgtttaa ttttacttct ccactacgtt 25740
 acgaaaatgg cagtcttaca ttcaactaca cttctcccat gacactaata aacaacagtc 25800
 ttgctattaa cgtcaatacc tccaaaggcc tcagtagtga caacggcaca ctcgctgtaa 25860
 atgttactcc agattttaga tttaacagct ctggtgcctt aacttttggc atacaaagtc 25920
 tatggacttt tccaacaaaa actcctaact gtaccgtggt taccgaaagt gactccctgc 25980
 tgagtctttg cttgactaaa tgcggagctc acgtacttgg aagcgtgagt ttaagcggag 26040
 tggcaggaac catgctaaaa atgaccacac cttctgttac cgttcagttt tcgtttgatg 26100

acagtggtaa	actaatattc	tctccacttg	cgaacaacac	ttgggggtgtt	cgacaaagcg	26160
agagtccgtt	gcccaccca	tccttcaacg	ctctcacgtt	tatgccaaac	agtaccattt	26220
attctagagg	agcaagtaac	gaacctcaaa	acaattatta	tgtccagacg	tatcttagag	26280
gcaacgtgcg	aaagccaatt	ctactaactg	ttacctacaa	ctcagttaat	tcaggatatt	26340
ccttaacttt	taaatgggat	gctgtcgcca	atgaaaaatt	tgccactcct	acatcttcgt	26400
tttgctatgt	tgagagcaa	taaaaccctg	ttaccccacc	gtctcgtttt	tttcagatga	26460
aacgagcgag	agttgatgaa	gacttcaacc	cagtgtaccc	ttatgacccc	ccatacgctc	26520
ccgtcatgcc	cttcattact	ccgcctttta	cctcctcgga	tggttgagc	gaaaaaccac	26580
ttggagtgtt	aagtttaaac	tacagggatc	ccattactac	acaaaatggg	tctctcacgt	26640
taaaactagg	aaacggcctc	actctaaaca	accagggaca	gttaacatca	actgctggcg	26700
aagtggagcc	tccgctcact	aatgctaaca	acaaacttgc	actagcctat	agcgaaccat	26760
tagcagtaaa	aagcaaccgc	ctaactctat	cacacaccgc	tccccttgct	atcgctaata	26820
attcttttagc	gttgcaagtt	tcagagccta	tttttgtaaa	tgacgatgac	aagctagccc	26880
tgagacagc	cgccccctt	gtaaccaacg	ctggcaccct	tcgcttacag	agcgctgccc	26940
ctttaggatt	ggttgaaaat	actcttaaac	tgctgttttc	taaacccttg	tatttgcaaa	27000
atgattttct	tgcatagacc	attgaacgcc	ccctggctgt	agcagccgca	ggtactctga	27060
ccctacaact	tactcctcca	ttaaagacta	acgatgacgg	gctaactacta	tccacagtcg	27120
agccattaac	tgtaaaaaac	ggaaacctag	gcttgcaaat	atcgcgccct	ttagtgtgtc	27180
aaaacaacgg	cctttcgctt	gctattaccc	ccccgctgcg	tttgtttaac	agcgaccccg	27240
ttcttggttt	gggcttcact	tttcccctag	ctgtcacaaa	caacctctc	tccttaaaaca	27300
tgggagacgg	agttaaactt	acctataata	aactaacagc	caatttggtt	agggatttac	27360
aatttgaaaa	cggtgagatt	gccgtaacgc	ttactgccga	attacctttg	caatacacta	27420
acaaacttca	actgaatatt	ggagctggcc	ttcgttacaa	tgagagccagc	agaaaactag	27480
atgtaaacat	taaccaaatt	aaaggcttaa	cttgggacaa	cgatgcagtt	attcccaaac	27540
taggatcggg	cttacaattt	gaccctaatt	gcaacatcgc	tgttatccct	gaaaccgtga	27600
agccgcaaac	gttatggagc	actgcagatc	cctcgcctaa	ctgctcagtg	taccaggact	27660
tggtatccag	gctgtggctc	gctcttggtt	aaagtggcga	catggtgcat	ggaagcattg	27720
ccctaaaagc	cctaaaaggg	acgttgctaa	atcctacagc	cagctacatt	tccattgtga	27780
tatatatttta	cagcaacgga	gtcaggcgta	ccaactatcc	aacgtttgac	aacgaaggca	27840
ccttagctaa	cagcgccact	tggttgatacc	gacaggggca	atctgctaac	actaatgtga	27900
ccaatgccac	tgaatttatg	cccagctcaa	gcaggtaccc	cgtgaataaa	ggagacaaca	27960
ttcaaaatca	atctttttca	tacacctgta	ttaaaggaga	ttttgctatg	cctgtcccgt	28020
tccgtgtaac	atataatcac	gccctggaag	gggtattccct	taagttcacc	tggtcgcttg	28080
tagccaatca	ggcctttgat	attccttgct	gttcattttc	atacatcaca	gaataaaaaa	28140

ccactttttc attttaattt ctttttattt tacacgaaca gtgagacttc ctccaccctt 28200
 ccatttgaca gcatacacca gcctctcccc cttcatagca gtaaactgtt gtgaatcagt 28260
 ccggtatttg ggagttaaaa tccaaacagt ctctttggtg atgaaacgtc gatcagtaat 28320
 ggacacaaat ccctgggaca ggttttccaa cgtttcggtg aaaaactgca caccgcccta 28380
 caaaacaaac aggttcaggc tctccacggg ttatctcccc gatcaaactc agacagggta 28440
 aaggtgcggt ggtgttcac taaaccacgc aggtggcgct gtctgaacct ctcggtgcga 28500
 ctctgtgag gctggtaaga agtttagattg tccagtagcc tcacagcatg tatcatcagt 28560
 ctacgagtgc gtctggcgca gcagcgcac tgaatctcac tgagattccg gcaagaatcg 28620
 cacaccatca caatcagggtt gttcatgatc ccatagctga acacgctcca gccaaagctc 28680
 attcgtctca acagcgccac cgcggtgtccg tccaacctta ctttaacata aatcagggtgt 28740
 ctgccgcgta caaacatgct acccacatac agaacttccc ggggcaggcc cctgttcacc 28800
 acctgtctgt accagggaaa cctcacattt atcagggagc catagatggc cattttaaac 28860
 caattagcta ataccgcccc accagctcta cactgaagag aaccgggaga gttacaatga 28920
 cagtgaataa tccatctctc ataaccctg atggtctgat gaaaatctag atctaacgtg 28980
 gcacaacaaa tacacacttt catatacatt ttcataacat gtttttccca ggccgttaaa 29040
 atacaatccc aatacacggg ccactcctgc agtacaataa agctaataca agatggtata 29100
 ctctcacct cactgacact gtgcatgttc atattttcac attctaagta ccgagagttc 29160
 tcctctacag cagcactgct gcggtcctca caaggtggta gctggtgatg attgtagggg 29220
 gccagtctgc agcgataccg tctgtcgcgt tgcacgtag accaggaacc gacgcacctc 29280
 ctcgtacttg tggtagcaga accacgtccg ctgccagcac gtctccacgt aacgccggtc 29340
 cctgcgtcgc tcacgtccc tcctcaatgc aaagtgaac cactcttgta atccacacag 29400
 atccctctcg gcctccgggg tgatgcacac ctcaaacta cagatgtctc ggtacagttc 29460
 caaacacgta gtgagggcga gttccaacca agacagacag cctgatctat cccgacacac 29520
 tggaggtgga ggaagacacg gaagaggcat gttattccaa gcgattcacc aacgggtcga 29580
 aatgaagatc ccgaagatga caacggctgc ctccggagcc ctgatggaat ttaacagcca 29640
 gatcaaacgt tatgcgattc tccaagctat cgatcgccgc ttccaaaaga gcctggaccc 29700
 gcactttcac aaacaccagc aaagcaaaag cactattatc aaactcttca atcatcaagc 29760
 tgcaggactg tacaatgcct aagtaatttt cgtttctcca ctcgcgatg atgtcgcggc 29820
 agatagtctg aaggttcac cctgacggg taaaaagctc cgaaagggcg ccctctacag 29880
 ccatgcgtag acacaccatc atgactgcaa gatatcgggc tcctgagaca cctgcagcag 29940
 atttaacaga tcaaggtcag gttgctctcc gcgatcacga atctccatcc gcaaggtcat 30000
 ttgcaaaaaa ttaaataaat ctatgccgac tagatctgtc aactccgcat taggaaccaa 30060
 atcaggtgtg gctacgcagc acaaaagtcc cagggtggt gccaaactca ctagaaccgc 30120
 tcccagtaaa caaaactgat gaatgggagt aacacagtgt aaaatgtgca accaaaaatc 30180
 actaaggtgc tcctttaaaa agtccagtag ttctatatcc agtccgtgca agtactgaag 30240

caactgtgcg ggaatatgca caacaaaaaa aatagggcg gctcagataca tgttgaccta 30300
 aaataaaaag aatcattaaa ctaaagaagc ttggcgaaac gtgggataaa tgacacgctc 30360
 cagcagcaga caggcaaccg gctgtccccg ggaaccgcg taaaattcat ccgaatgatt 30420
 aaaaagaaca acagaaactt cccaccatgt actcggtttg atctcctgag cacacagcaa 30480
 tccccccctc acattcatgt ccgccacaga aaaaaaacgt cccagatacc cagcggggat 30540
 atccaacgac agctgcaaag acagcaaaac aatccctctg ggagcgatca caaaatcctc 30600
 cgggtgaaaaa agcacatata tattagaata accctgttgc tggggcaaaa aggcccggcg 30660
 tcccagcaaa tgcacataaa tatgttcatc agccattgcc ccgtcttacc gcgtaatcag 30720
 ccacgaaaaa atcgagctaa aattcaccca acagcctata gctatatata cactccgccc 30780
 aatgacgcta ataccgcacc acccacgacc aaagttcacc cacaccaca aaaccgcga 30840
 aaatccagcg ccgtcagcac ttccgcaatt tcagtctcac aacgtcactt ccgcgcgcct 30900
 tttcacattc ccacacacac ccgcgccctt cgcgccgccc tcgcgccacc ccgcgtcacc 30960
 gcacgtcacc ccggccccgc ctgcgtcctc cccgctcatt atcatattgg cacgtttcca 31020
 gaataaggta tattattgat gatg 31044

<210> 6
 <211> 34115
 <212> DNA
 <213> simian adenovirus SV-39

<400> 6
 catcatcaat ataacaccgc aagatggcga ccgagttaac atgcaaata ggtgggcgga 60
 gttacgcgac ctttgtcttg ggaacgcgga agtgggcgcg gcgggtttcg gggaggagcg 120
 cggggcgggg cgggcggtgc gcgcggcggt gacgcgccgg ggaccggaa attgagtagt 180
 ttttattcat ttgcaagtt tttctgtaca ttttggcgcg aaaactgaaa cgaggaagtg 240
 aaaagtgaaa aatgccgagg tagtcaccgg gtggagatct gacctttgcc gtgtggagtt 300
 taccgcgtga cgtgtgggtt tcggtctcta ttttttact gtggttttcc gggtagcgtc 360
 aaaggtcccc attttatgac tccacgtcag ctgatcgta gggattttaa tgcgcctcag 420
 accgtcaaga ggccactctt gagtgcgggc gagaagagtt ttctcctccg cgttccgcca 480
 actgtgaaaa aatgaggaac ttcttgctat ctccggggct gccagcgacc gtagccgccg 540
 agctgttgga ggacattgtt accggagctc tgggagacga tcctcaggtg atttctcact 600
 tttgtgaaga ttttagtctt catgatctct atgatattga tccgggtgtt gaggggcaag 660
 aggatgaatg gctggagtct gtggatgggt tttttccgga cgctatgctg ctagaggctg 720
 atttgccacc acctcacaac tctcacactg agcccgagtc agctgctatt cctgaattgt 780
 catcaggtga acttgacttg gcttgttacg agactatgcc tccggagtcg gatgaggagg 840
 acagcgggat cagcgatccc acggctttta tgggtctctaa ggcgattgct atactaaaag 900
 aagatgatga tggcgatgat ggatttcgac tggacgctcc ggcggtgccg gggagagact 960
 gtaagtcctg tgaataccac cgggatcgta ccggagacct gtctatgttg tgttctctgt 1020

gttatctccg	tcttaacgct	gcttttgtct	acagtaagtg	ttttgtgctt	ttttaccctg	1080
tggctttgtt	gagtttattt	ttttctgtgt	ctcatagggg	gttgtttatt	ataggtcctg	1140
tttcagatgt	ggaggaacct	gatagtacta	ctggaaatga	ggaggaaaag	ccctccccgc	1200
cgaaactaac	tcagcgctgc	agacctata	ttttgagacc	ctcggcccag	cgtgtgtcat	1260
cccggaaacg	tgctgctgtt	aattgcatag	aagatttatt	ggaagagccc	actgaacctt	1320
tggacttgct	cttaaagcga	ccccgcccgc	agtagggcgc	ggtgccagtt	ttttctctct	1380
agcttccggg	tgactcagtg	caataaaaa	tttcttgcca	acaggtgtat	gtgtttactt	1440
tacgggcggg	aagggattag	gggagtataa	agctggaggg	gaaaaatctg	aggctgtcag	1500
atcgagttag	aagttccatg	gacttgtagc	agagcctaga	gaatctaagt	tctttgtagc	1560
gtttgctgga	ggaggcctcc	gacagaacct	cttacatttg	gaggtttctg	ttcggttccc	1620
ctctgagtcg	ctttttgcac	cgggtgaagc	gagagcacct	gacggaattt	gatgggcttt	1680
tagagcagct	gcctggactg	tttgattctt	tgaatctcgg	ccaccggacg	ctgctagagg	1740
agaggctttt	tccacaattg	gacttttctt	ctccaggccg	tctgtgttca	gcgcttgctt	1800
ttgctgtaca	tctgttgtag	agatggaacg	agcagacgca	gctcagcccc	ggttacactc	1860
tggacttcct	gacgctatgc	ctatggaagt	tcggaatcag	gagggggagg	aagctgtacg	1920
ggcgcttggt	ggagaggcat	ccgtctctgc	gccagcagcg	tctgcaagct	caagtgtctg	1980
tgaggcggga	ggatctggaa	gccatttcgg	aggaggagag	cggcatggaa	gagaagaatc	2040
cgagagcggg	gctggaccct	ccggcggagg	agtagggggg	ataccggacc	cttttcctga	2100
gttggttttg	ggggcggttg	ggggcgcttc	tgtggtacgt	gaggatgaag	aggggcgcca	2160
acgcggtcag	aagaggggag	atgttgagtc	ctcgactttc	ttggctgatg	taaccgtggc	2220
cctgatggcg	aaaaacaggc	tggaggtggt	gtggtacccg	gaagtatggg	aggactttga	2280
gaagggggac	ttgcacctgc	tggaaaaata	taactttgag	caggtgaaaa	catactggat	2340
gaacccggat	gaggactggg	aggtggtttt	gaaccgatac	ggcaaggtag	ctctgcgtcc	2400
cgactgtcgc	taccaggttc	gcgacaagg	ggtcctgcga	cgcaacgtgt	acctgttggg	2460
caacggcgcc	accgtggaga	tgggtggacc	cagaaggggt	ggttttgtgg	ccaatatgca	2520
agaaatgtgc	cctgggggtg	tgggcttgct	tggggtgact	tttcatagt	tgaggtttag	2580
cggtagcaat	tttgggggtg	tggttattac	cgcgaacact	cctgtggtcc	tgcataattg	2640
ctactttttt	ggcttcagca	acacctgtgt	ggaaatgagg	gtgggaggca	aagtgcgcgg	2700
gtgttccttt	tacgcttgct	ggaagggggg	ggtgagccag	ggtaaggcta	aagtgtctgt	2760
tcacaagtgt	atgttgtaga	gatgcacctt	gggcatttcc	agtgagggct	tcctccacgc	2820
cagcgacaac	gtggcttctg	acaacggctg	cgcctttctt	atcaaggagg	ggggctcgcat	2880
ctgtcacaac	atgatatgag	gccctgggga	tgtcccccca	aagccttacc	agatggttac	2940
ctgcacagat	ggcaagggtg	gcatgctcaa	gcctgtgcac	attgtggggc	accggcgcca	3000
ccgctggcca	gagtttgaac	acaatgtgat	gacccgctgt	agcttgtacc	tgggaggcag	3060

gcgaggagtt	ttcttgccca	gacagtgtaa	cctggcccac	tgcaacgtga	tcatggaaca	3120
atccgccgct	accaggttt	gctttggagg	aatatttgat	ataagcatgg	tggtgtataa	3180
gatcctgcgc	tacgacgact	gtcgggctcg	tactcgaacc	tgcgactgcg	gagcctctca	3240
cctgtgtaac	ctgactgtga	tggggatggt	gactgaggag	gtgcgactgg	accactgtca	3300
gcactcttgc	ctgcgggagg	agttttcttc	ctcggacgag	gaggactagg	taggtggttg	3360
gggcgtggcc	agcgagaggg	tgggctataa	aggggaggtg	tcggctgacg	ctgtcttctg	3420
tttttcaggt	accatgagcg	gatcaagcag	ccagaccgcg	ctgagcttcg	acggggccgt	3480
gtacagcccc	tttctgacgg	ggcgcttgcc	tgcttgggcc	ggagtgcgtc	agaatgttac	3540
cggttcgacc	gtggacggac	gtcccgtgga	tccatctaac	gctgcttcta	tgcgctacgc	3600
tactatcagc	acatctactc	tggacagcgc	cgctgccgcc	gcagccgcca	cctcagccgc	3660
tctctccgcc	gccaagatca	tggctattaa	cccaagcctt	tacagccctg	tatccgtgga	3720
cacctcagcc	ctggagcttt	accggcgaga	tctagctcaa	gtggtggacc	aactcgcagc	3780
cgtgagccaa	cagttgcagc	tggtgtcgac	ccgagtggag	caactttccc	gccctcccca	3840
gtaaccgcaa	aaattcaata	aacagaattt	aataaacagc	acttgagaaa	agtttaaact	3900
tgtggttgac	tttattcctg	gatagctggg	gggagggaac	ggcgggaacg	gtaagacctg	3960
gtccatcggt	cccggtcggt	gagaacacgg	tggatttttt	ccaagacctg	atagaggtgg	4020
gtctgaacgt	tgagatacat	gggcatgagc	ccgtctcggg	ggtggaggta	ggccactgc	4080
agggcctcgt	tttcaggggt	ggtgttgtaa	atgatccagt	cgtaggcccc	ccgctgggcg	4140
tggtgctgga	agatgtcctt	cagcagcaag	ctgatggcaa	cgggaagacc	cttggtgtag	4200
gtgttgacaa	agcggttgag	ttgggagggg	tgcatgcggg	gactgatgag	gtgcattttg	4260
gcctggatct	tgaggttggc	tatgttgccg	cccagatcgc	gcctgggatt	catgttatgc	4320
aagaccacca	gcaccgagta	accggtgcag	cgggggaatt	tgctgtgcag	cttggaaggg	4380
aaagcgtgga	agaatttgga	gaccctcggg	tgccgccta	ggttttccat	gcactcatcc	4440
atgatgatgg	cgatggggcc	ccgggaggca	gcctgggcaa	aaacgttgcg	ggggtccgtg	4500
acatcgtagt	tgtggtcctg	ggtgagttca	tcataggaca	ttttgacaaa	gcgcgggcag	4560
agggtccag	actggggaat	gatggttcca	tccggtccgg	gggcgtagtt	gccctcgag	4620
atttgcattt	cccaggcttt	gatttcagag	ggagggatca	tgtcaacctg	gggggcgatg	4680
aaaaaatgg	tctctggggc	gggggtgatg	agctgggtgg	aaagcagggt	gcgcaagagc	4740
tgtgacttgc	cgagccgggt	gggcccgtag	atgacagcta	tgacgggttg	cagggtgtag	4800
tttagagagc	tacaactgcc	atcatccttc	aaaagcgggg	ccacactgtt	taaaagtctt	4860
ctaacatgta	agttttcccg	cactaagtcc	tgaggagac	gtgaccctcc	tagggagaga	4920
agttcaggaa	gcgaagcaaa	gtttttaagt	ggcttgaggc	catcggccaa	gggcaagttc	4980
ctgagagttt	gactgagcag	ttccagccgg	tcccagagct	cggttacgtg	ctctacggca	5040
tctcgatcca	gcagacctcc	tcgtttcggg	ggttggggcg	gctctggctg	tagggaatga	5100
ggcgggtgggc	gtccagctgg	gccatggtgc	ggtccctcca	tgggcgcagg	gttctcttca	5160

gggtggtctc	ggtcacggtg	aatgggtggg	ccccgggctg	ggcgtggcc	agggtgcgct	5220
tgaggctgag	gcggtggtg	gcgaaccgtt	gcttttcgtc	tccctgcaag	tcagccaaat	5280
agcaacggac	catgagctca	tagtccaggc	tctctgcggc	atgtcctttg	gcgcgaagct	5340
tgcttttgga	aacgtgccc	cagtttgagc	agagcaagca	ttttagcgcg	tagagttttg	5400
gcgccaagaa	cacggattcc	ggggaataag	catccccacc	gcagttggag	caaacggttt	5460
cgcattccac	cagccaggtc	agctgaggat	cttttggtc	aaaaaccaag	cgcccgcgct	5520
tttttttgat	gcgcttccta	cctcggtct	ccatgaggcg	gtgcccgcgt	tcggtgacga	5580
agaggctgtc	ggtgtctccg	tagacggagg	tcagggcgcg	ctcctccagg	ggggtcccgc	5640
ggtcctcggc	gtagagaaac	tcgcaccact	ctgacataaa	cgcccgggtc	caggctagga	5700
cgaatgaggc	gatgtgggaa	gggtaccggt	cgttatcgat	gagggggctg	gttttttcca	5760
aggtgtgcag	gcacatgtcc	ccctcgtccg	cttccaaaaa	tgtgattggc	ttgtaggtgt	5820
aagtcacgtg	atcctgtcct	tccgcggggg	tataaaagg	ggcgtttccc	ccctcctcgt	5880
cactctcttc	cggttcgctg	tcgccaaagg	ccagctgttg	gggtacgtaa	acgcgggtga	5940
aggcgggcat	gacctgtgcg	ctgaggttgt	cagtttctat	atacgaggaa	gatttgatgg	6000
cgagcgcccc	cgtggagatg	cccttgaggt	gctcggggcc	catttggtca	gaaaacacaa	6060
tctgtcgggt	atcaagcttg	gtggcaaaag	acccgtagag	ggcgttgag	agcaacttgg	6120
cgatggagcg	ctgggttttg	tttttttccc	ggtcggcttt	ttccttgcc	gcgatgttga	6180
gctggacgta	ctccctggcc	acgcacttcc	agccgggaaa	aacggccgtg	cgctcgtccg	6240
gcaccagcct	cacgtcccat	ccgcggttgt	gcagggtgat	gacgtcgatg	ctggtggcca	6300
cctctccgcg	caggggctcg	ttggtccagc	agaggcgacc	gcccttgca	gagcagaagg	6360
ggggcagggg	gtcaagcagg	cgctcgtccg	gggggtcggc	gtcgatggta	aagatggcgg	6420
gcagcaggtg	tttgtcaaag	taatcgatct	gatgcccggg	gcaacgcagg	gcggtttccc	6480
agtcccgcac	cgccaaggcg	cgctcgtatg	gactgagggg	ggcgccccag	ggcatgggat	6540
gcgtcagggc	cgaggcgtag	atgccgcaga	tgtcatagac	gtaaaggggc	tcctccagga	6600
cgccgaggtg	ggtggggtag	cagcgcccc	cgcggtgct	ggcccgtacg	tagtcgtaga	6660
gctcgtgcga	gggggcccaga	aggtggcggc	tgagggtgagc	gcgctggggc	ttttcatctc	6720
ggaagaggat	ctgcctgaag	atggcgtggg	agttggagga	gatggtgggc	cgctgaaaaa	6780
tgttgaagcg	ggcgtcgggc	agaccacagg	cctcgccgat	aaagtgggcg	taggactctt	6840
gcagcttttc	caccaggag	gcggtgacca	gcacgtccag	agcgcagtag	tccagggttt	6900
cccgcacgat	gtcataatgc	tcttcctttt	tttccttcca	gaggctctcg	ttgaagagat	6960
actcttcg	gtctttccag	tactcttgga	gaggaaaccc	gttttcgtct	ccacggtgaa	7020
agcccaacat	gtaaaactgg	ttgacggcct	gatagggaca	gcaccccttc	tccacgggca	7080
gcgagtaggc	cagggcgggc	ttgcgcaggg	aggtgtgagt	cagggcaaag	gtgtcgcgga	7140
ccataacttt	tacaaactgg	tacttaaagt	cccggctcgtc	gcacatgcct	cgctcccagt	7200

ctgagtagtc	tgtgcgcttt	ttgtgcttgg	ggtaggcag	ggagtaggtg	acgtcgtaa	7260
agaggatttt	gccacatctg	ggcataaagt	tgcgagagat	tctgaagggg	ccgggcacct	7320
ccgagcgggt	gttgatgact	tgggcagcca	ggagaatttc	gtcgaagccg	ttgatgttgt	7380
gccccacgac	gtagaactct	atgaaacgcg	gagcgccgcg	cagcaggggg	cacttttcaa	7440
gttgctgaa	agtaagttcc	cgcggtctga	cgccgtgttc	cgtgcggtc	cagtcctcca	7500
ccgggtttcg	ctccacaaaa	tcctgccaga	tgtggtcgac	tagcaagagc	tgcagtcggt	7560
cgcgaaattc	gcggaatttt	ctgccgatgg	cttgcttctg	ggggttcaag	caaaaaaagg	7620
tgtctgcgtg	gtcgcgccag	gcgtcccagc	cgagctcgcg	agccagattc	agggccagca	7680
gcaccagagc	cggctcaccg	gtgattttca	tgacgaggag	aaagggcacc	agctgttttc	7740
cgaacgcgcc	catccaggtg	taggtctcca	cgtcgtaggt	gagaaacaga	cgttcggtcc	7800
gcgggtgcga	tcccaggggg	aaaaacttga	tgggctgcc	ccattgggag	ctctgggcgt	7860
ggatgtgatg	gaagtaaaag	tcccggcggc	gcgtggaaca	ttcgtgctgg	tttttgtaaa	7920
agcggccgca	gtggtcgcag	cgcgagacgg	agtgaaggct	gtgaatcagg	tgaatcttgc	7980
gtcgtgagg	gggccccaga	gccaaaaagc	ggagcgggaa	cgaccgcgcg	gccacttcgg	8040
cgtccgcagg	caagatggat	gagggttcca	ccgttccccg	cccgcggacc	gaccagactt	8100
ccgccagctg	cggcttcagt	tcttgacca	gctctcgcag	cgtttcgtcg	ctgggcgaat	8160
cgtgaatacg	gaagttgtcg	ggtagaggcg	ggaggcggtg	gacttccagg	aggtgtgtga	8220
gggccggcag	gagatgcagg	tggtacttga	tttccacagg	atgacggtcg	cgggcgtcca	8280
aggcgaagag	atgaccgtgg	ggccgcggcg	ccaccagcgt	tccgcggggg	gtctttatcg	8340
gcggcgggga	cgggctcccc	gcggcagcgg	cggctcggga	cccgcgggca	agtcgggcag	8400
cggcacgtcg	gcgtggagct	cgggcagggg	ctggtgctgc	gcgcggagct	gactggcaaa	8460
ggctatcacc	cggcgattga	cgtcctggat	ccggcggcgc	tgctgaaga	ccaccggacc	8520
cgtggtcttg	aacctgaaag	agagttcgac	agaatcaatc	tcggcatcgt	taaccgcggc	8580
ctggcgcagg	atttcggcca	cgtccccgga	gttgcttga	tacgcgattt	ctgcatgaa	8640
ctggtcgatt	tcctcttcct	gcaagtctcc	gtgaccggcg	cgttcgacgg	tggccgcgag	8700
atcgttggag	atgcggccca	tgagctggga	aaaggcattg	atgccgacct	cgttccacac	8760
tcggctgtac	accacctctc	cgtgaacgtc	gcgggcgcgc	atcaccacct	gggcgagatt	8820
gagttccacg	tggcgggcca	aaaccggata	gtttcgagg	cgctgataca	gatagttgag	8880
gggtggtggcg	gcgtgctcgg	ccacaaaaaa	atacatgatc	cagcggcgga	gggtcagctc	8940
gttgatgtcg	cccagcgcct	ccaggcgttc	catggcctcg	taaaagtcca	cggcaaagtt	9000
gaaaaattgg	ctgttcctgg	ccgagaccgt	gagctcttct	tccaagagcc	gaatgagatc	9060
cggcacgggtg	gccctgactt	cgcgttcgaa	agccccgggt	gcctcctcca	cctcttcctc	9120
ctcgacttct	tcgaccgctt	cgggcacctc	ctcttcctcg	accaccacct	caggcggggc	9180
tcggcggcgc	cggcggcgga	cgggcaggcg	gtcgacgaaa	cgctcgatca	tttccccctt	9240
ccgtcgacgc	atggtctcgg	tgacggcgcg	accctgttctg	cgaggacgca	gggtgaaggc	9300

gccgccgccg agcggaggta acagggagat cggggggcgg tcgtggggga gactgacggc 9360
 gctaactatg catctgatca atgtttgcgt agtgacctcg ggtcggagcg agctcagcgc 9420
 ttgaaaatcc acgggatcgg aaaaccgttc caggaacgcg tctagccaat cacagtcgca 9480
 aggtaagctg aggaccgtct cgggggcttg tctgttctgt cttcccgcgg tgggtgctgct 9540
 gatgaggtag ttgaagtagg cgctcttgag gcggcggtat gtggacagga gaaccacgtc 9600
 tttgcgcccc gcttgctgta tccgcaggcg gtcggccatg cccacactt ctccttgaca 9660
 gcggcgaggg tccttgtagt attcttgcgt cagcctttcc acgggcacct cgtcttcttc 9720
 ttccgctcgg ccggacgaga gccgcgtcag gccgtacctg cgctgccccct gtggttgagg 9780
 cagggccagg tcggccacga cgcgctcggc cagcacggcc tgctggatgc ggggtgagggt 9840
 gtcctgaaag tcgtcgagat ccacaaagcg gtggtacgcg ccagtgttga tgggtgtaggt 9900
 gcagttgctc atgacggacc agtttacggt ctgggtgcca tggcccacgg tttccaggta 9960
 gcggagacgc gagtagggcc gcgtctcgaa gatgtagtcg ttgcagggtc gcagcaggta 10020
 ctggtagccc accagcagat gcggcggcgg ctggcggtag aggggcccacc gctgggtggc 10080
 gggggcggtg ggggcgagat cttccaacat gaggcggtga tagccgtaga tgtagcgca 10140
 catccaagtg atgccgctgg ccgtgggtgct ggcgcgggcg tagtcgcgaa cgcggttcca 10200
 gatgtttcgc agcggctgga agtactcgat ggtggggcga ctctgccccg tgaggcgggc 10260
 gcagtcggcg atgctctacg gggaaaaaga agggccagt aacaaccgcc ttccgtagcc 10320
 ggaggagaac gcaagggggg caaagaccac cgaggctcgg gttcgaaacc cgggtggcgg 10380
 cccgaatacg gagggcggtt ttttgctttt ttctcagatg catcccgtgc tgcggcagat 10440
 gcgtccgaac gcggggtccc agtccccggc ggtgcctgcg gccgtgacgg cggttctac 10500
 ggccacgtcg cgctccaccc cgcctaccac ggcccaggcg gcggtggctc tgcgcggcgc 10560
 aggggaaccc gaagcagagg cgggtgttga cgtggaggag ggccaggggt tggctcggct 10620
 gggggccctg agtcccagc ggcacccgcg cgtggctctg aagcgcgacg cggcgaggc 10680
 gtacgtgccg cggagcaatc tgtttcgcga ccgcagcggc gaggaggccg aggagatgcg 10740
 agacttgctg tttcgggcgg ggaggaggt gcgtcacggg ctggaccggc agagggttct 10800
 gagagaggag gactttgagg cggacgagcg cacgggggtg agtcccgcgc gggctcacgt 10860
 ggcgccgcc aacctggtga gcgcgtacga gcagacggtc aaggaggaga tgaacttcca 10920
 gaagagcttc aatcatcacg tgcgcacgct gattgcgcgc gaagagggtg ccatcggcct 10980
 catgcatctg tgggattttg tggaggcgta cgttcagaac cccagcagca agccgctgac 11040
 ggctcagctg ttcctcatcg tgcaacatag tcgagacaac gaaacgttca gggaggccat 11100
 gctgaacatt gcagagcctg aggggcgctg gctcttggtat ctattaaca tcttgagag 11160
 tatcgtagtg caggagcgt cgctgagcct ggccgacaag gtggctgcca tcaactacag 11220
 catgctgtcg ctgggcaaatt ttacgccc caagatctac aagtctccgt tcgtcccat 11280
 agacaaggag gtgaagatag acagctttta catgcgcatg gcgctcaagg tgctgactct 11340

aagcgacgac ctgggggtgt accgcaacga cgcatacac aaggcggtga gcgccagccg 11400
ccggcgcgag ctgagcgacc gcgagctttt gcacagcctg catcgggctg tgactggtgc 11460
cggcagcgcc gaggcggccg agtactttga cgccggagcg gacttgcgct ggcagccatc 11520
ccgacgcgcg ctggaggcgg ctggcgctcg ggagtaggg gtcgaggacg acgatgaagc 11580
ggacgacgag ttgggcattg acttgtagcc gtttttcgtt agatatgtcg gcgaacgagc 11640
cgtctgcggc cgccatggtg acggcgggcg gcgcgcccc ggacccggcc acgcgcgcgg 11700
cgctgcagag tcagccttcc ggagtgacgc ccgcggacga ctggtccgag gccatgcgtc 11760
gcatcctggc gctgacggcg cgcaaccccc aggccttttcg gcagcagccg caggcaaacc 11820
ggtttgcggc ctttttgaa gcggtggtgc cctccagacc caaccacc caccgaaaagg 11880
tgctggccat cgtcaacgcc ctggcggaga ccaaggccat ccgccagac gaggccgggc 11940
aggtttacaa cgcgctgcta gaaagggtgg gacgctacaa cagctccaac gtgcagacca 12000
atctggaccg cttggtgacg gacgtgaagg aggccgtagc ccagcgagag cggtttttca 12060
aggaagccaa tctgggctcg ctggtggccc tcaacgcctt cctgagcacg ctgccggcga 12120
acgtgccccg cggtcaggag gactacgtga actttctgag cgccctccgc ctgatggtgg 12180
ccgaggtgcc gcagagcgag gtgtaccagt ctggcccaa ctactacttc cagacctccc 12240
ggcagggcct gcagacggtg aacctgacgc aggcctttca gaacctgcag ggcctttggg 12300
gggtgcgcgc tccgctgggc gaccgcagca cgggtgtccag cctgctgacc cccaatgccc 12360
ggctgctctt gcttctcatt gctccgttca ccgacagcgg ttccatcagc cgcgactctt 12420
acctgggaca cctgctcacc ctgtaccggg aggccatcgg gcaggcgcgg gtggacgagc 12480
agacgtacca ggaaatcacc agcgtgagcc gcgcgctggg gcaggaggac acgggcagct 12540
tgagggcgac tctgaacttc ctgctgacca accggcgga gcgcctacct cccagtagc 12600
cgctgaacgc ggaggaggag cgcacctgc gtttcgtgca gcagagcacc gcgctgtact 12660
tgatgcggga aggcgcctct cccagcgctt cgctggacat gacggcgggc aacatggagc 12720
catcgttcta cgccgccaac cgtcccttcg tcaaccggct aatggactat ttgcatcggg 12780
cggcgccct gaaccggaa tactttacta acgtcatcct gaacgaccgt tggctgccac 12840
ctcccgctt ctacacggg gagttcgacc tcccggaggc caacgacggt ttcattgtgg 12900
acgacgtgga cagcgtgttc ctgcccggca agaaggaggc gggtgactct cagagccacc 12960
gcgcgagcct cgagacctg gggcgaccg gggcgcgctc tccgctgcct cgcctgccga 13020
gcgccagcag cgccagcgtg gggcggtga gccgtccgc cctcagcggg gaggaggact 13080
ggtggaacga tccgctgctc cgtccggccc gcaacaaaaa cttcccaac aacgggatag 13140
aggatttggg agacaaaatg aaccgttgga agacgtatgc ccaggagcat cgggagtggc 13200
aggcgaggca acccatgggc cctgttctgc cgccctctcg gcgcccgcgc agggacgaag 13260
acgccgacga ttcagccgat gacagcagcg tgttgatct gggcgggagc ggaaccct 13320
ttgccacct gcaacctgc ggcgtgggtc ggcggtggcg ctaggaaaaa aaattattaa 13380
aagcacttac cagagccatg gtaagaagag caacaaagg gtgtcctgct ttcttcccgg 13440

tagcaaaatg cgtcgggagg tggcagttcc ctccgcggca atggcgtag gcccgcccc 13500
 ttcttacgaa agcgtgatgg cagcggccac cctgcaagcg ccgttgaga atccttacgt 13560
 gccgccgca tacctggagc ctacgggagg gagaaacagc attcgttact cggagctgac 13620
 gccctgtac gacaccacc gcctgtacct ggtggacaac aagtcagcag atatcgccac 13680
 cttgaactac cagaacgacc acagcaactt tctcacgtcc gtggtgcaga acagcgacta 13740
 cacgcccgc gaagcgagca cgcagaccat taacttggac gaccgctcgc gctggggagg 13800
 ggacttgaaa accattctgc aactaacat gcccaacgtg aacgagttca tgtttaccaa 13860
 ctcgttcagg gctaaactta tggtaggca caggggcagc aaggaccagg tttatgagtg 13920
 ggtgcagctg acgctgagg aggggaactt ttcagagatt atgaccatag acctgatgaa 13980
 caacgccatt atcgaccact acctggagg agccagacag cagggggtga aagaaagcga 14040
 gatcggcgtc aagtttgaca cgcgaactt tcgtctgggc tgggaccagg agacggggct 14100
 tgtgatgccg ggggtgtaca cgaacgaagc tttccatccc gacgtggtcc tcttgccggg 14160
 ctgaggggtg gactttacct acagccggtt aaacaacctg ctaggcatac gcaagagaat 14220
 gccctttcag gaagggttt agatcctgta cagggacctg gagggcggtg acatcccggc 14280
 cctgctggac gtgcccggc acgaggagag catcgccaac gcaagggagg cggcgatcag 14340
 gggcgataat ttcgcccgc agccccaggc ggctccaacc ataaaaccg ttttggaaga 14400
 ctcaaaggc cggagctaca acgtaatagc caacaccaac aacacggctt acaggagctg 14460
 gtatctggct tataactacg gcgacccgga gaagggggtt agggcctgga ccctgctcac 14520
 cactccggac gtgacgtgcg gttcagagca ggtctactgg tcgctgcctg acatgtacgt 14580
 ggaccctgtg acgtttcgt ccacgcagca agtttagcaac taccagtggt tgggagcggg 14640
 gcttatgccg attcacagca agagctttta caacgagcag gccgtctact cacagctcat 14700
 tcgtcagacc accgcccta cgcacgtttt caaccgctt cccgagaacc aaatcctagt 14760
 gcgacctca gcgcccacca tcaccaccgt cagcgagaac gtgcccgtc taaccgatca 14820
 cgggacgctg cctttgcaga acagcatccg cggagttcag cgagttacca tcacggacgc 14880
 ccgtcgtcgg acctgtccct acgtctacaa agccttgagg atcgtggccc cgcgcgtcct 14940
 gtcgagtcgc actttctaga tgtccatcct catctctccc agcaacaata ccggttgggg 15000
 tctgggagtg accaaaatgt acggaggcgc caaacgacgg tccccacaac atcccgtagc 15060
 agtgcgaggc cacttttagag ccccatgggg gtcgcacacg cgggggagca ccggccgaac 15120
 caccgtcagc gacgtgatcg atagcgtggg ggccgacgcc cgcaactacc agcccgctcg 15180
 atccacgggtg gacgaagtca tcgacggcgt ggtggccgac gccagggcct acgcccgcag 15240
 aaagtctcgt ctgcccgcg gccgttcgtt aaagcgcccc acggccgcca tgaaagccgc 15300
 tcgtctcttg ctgctcgcg cacgtatcgt gggtagccgc gccgagac gcgcagccgc 15360
 caacgccgcg gccggccgag tgcgcccgc ggccgcccag caggccgccc ccgcatctc 15420
 cagtctatcc gcccccgac gcgggaatgt gtactgggtc agggactcgg ccaccggcgt 15480

gcgagttccc gtgagaaccc gtcctcctcg tccctgaata aaaagttcta agcccaatcg 15540
 gtgttccggt gtgtgttcag ctggtcatga ccaaacgcaa gtttaagag gagctgctgc 15600
 aagcgctggt ccccgaaatc tatgcgccgg cgccggacgt gaaaccgct cgctgaaac 15660
 gcgtgaagaa gcaggaaaag ctagagacaa aagaggaggc ggtggcggtg ggagacgggg 15720
 aggtggagtt tgtgctctcg ttcgctccgc gtcggcgagt gaattggaag gggcgcaagg 15780
 tgcaacgggt gctgctccc ggcacggtgg tgtctttcac cccgggtgaa aaatccgcct 15840
 ggaagggcat aaagcgctg tacgatgagg tgtacgggga cgaagacatt ctggagcagg 15900
 cgctggatag aagcggggag tttgcttac gcaagagggc gaggacgggc gagatcgcca 15960
 tcccgctgga cacttccaac cccaccccca gtctgaaacc cgtgacgctg caacagggtgt 16020
 tgccgggtgag cgccccctcg cgacgcggca taaaacgcga gggcggcgag ctgcagccca 16080
 ccatgcagct cctggttccc aagaggcaga aactagagga cgtactggac atgataaaaa 16140
 tggagcccga cgtgcagccc gatattaaaa tccgtcccat caaagaagtg gcgccgggaa 16200
 tgggcgtgca gaccgtggac atccagattc ccatgaccag cgccgcacag gcggtagagg 16260
 ccatgcagac cgacgtgggg atgatgacgg acctgcccgc agctgctgcc gccgtggcca 16320
 gcgccgcgac gcaaacggaa gccggcatgc agaccgacc gtggacggag gcgccgctgc 16380
 agccggccag aagacgcgtc agacggacgt acggccccgt ttctggcata atgccggagt 16440
 acgcgctgca tccttccatc atccccaccc ccggctaccg ggggcgcacc taccgtccgc 16500
 gacgcagcac cactcgccgc cgtcgccgca cggcacgagt cgccaccgcc agagtgcagc 16560
 gcgtaacgac acgtcgccgc cgccgcttga ccctgcccgt ggtgctgtac catcccagca 16620
 ttctttaaaa aaccgctcct acgttgacga tgggcaagct tacttgtcga ctccgtatgg 16680
 ccgtgcccgg ctaccgagga agatcccgcc gacgacggac tttgggaggc agcggtttgc 16740
 gccgcgctc ggcggttcac cggcgctca agggaggcat tctgccggcc ctgatcccca 16800
 taatcgccgc agccatcggg gccattcccc gaatcgccag cgtagcggtg caggctagcc 16860
 agcgccactg attttactaa ccctgtcggc cgcgccgtct ctttcggcag actcaacgcc 16920
 cagcatggaa gacatcaatt tctctctctt ggccccgcgg cacggcacgc ggccgtatat 16980
 ggggacgtgg agcgagatcg gcacgaacca gatgaacggg ggcgctttca attggagcgg 17040
 tgtgtggagc ggcttgaaaa atttcggttc cactctgaaa acttacggca accgggtgtg 17100
 gaactccagc acggggcaga tgcagggga caagctaaag gacacgcagt ttcagcaaaa 17160
 ggtggtggac ggcacgctt cgggcctcaa cggcgccgtc gacctggcca accaggccat 17220
 tcaaaaggaa attaacagcc gcctggagcc gcggccgcag gtggaggaga acctgcccc 17280
 tctggaggcg ctgccccca agggagagaa gcgcccgcgg ccgacatgg aggagacgct 17340
 agttactaag agcgaggagc cgccatcata cgaggaggcg gtgggtagct cgcagctgcc 17400
 gtccctcacg ctgaagccca ccacctatcc catgaccaag cccatcgctt ccatggcgcg 17460
 ccccggtgga gtcgacccgc ccatcgacgc ggtggccact ttggacctgc cgcgccccga 17520
 acccggaac cgctgcctc ccgtcccat cgctccgccc gtttctcgcc ccgcatccg 17580

ccccgtcgcc gtggccactc cccgctatcc gagccgcaac gccaaactggc agaccaccct 17640
caacagtatt gtcggactgg ggggtgaagtc tctgaagcgc cgtcgctgtt tttaaagcac 17700
aatttattaa acgagtagcc ctgtcttaat ccatcgttgt atgtgtgcct atatcacgcg 17760
ttcagagcct gaccgtccgt caagatggcc actccgtcga tgatgccgca gtggtcgtac 17820
atgcacatcg ccgggcagga cgcctcggag tacctgagcc cgggtctggt gcagtttgcc 17880
cgtgcgacgg aaacctactt ctactgggc aacaagttca ggaacccac cgtggcgccc 17940
accacgacg tcaccaccga tcgggtccag cgactgacaa tccgcttcgt ccccgtaggac 18000
aaggaagaca ccgcttactc ctacaaaacc cgcttcacgc tggccgtggg cgacaaccgg 18060
gtgctagaca tggccagtac ctactttgac atccgcggcg tgatcgaccg cggacctagc 18120
ttcaagcctt actccggcac ggcttacaac tctactggctc ccaaaggggc gcccaacaac 18180
agccaatgga acgccacaga taacgggaac aagccagtgt gttttgctca ggcagctttt 18240
ataggtcaaa gcattacaaa agacggagtg caaatacaga actcagaaaa tcaacaggct 18300
gctgccgaca aaacttacca accagagcct caaattggag tttccacctg ggataccaac 18360
gttaccagta acgctgccgg acgagtgtta aaagccacca ctcccatgct gccatgttac 18420
ggttcatatg ccaatccac taatccaaac gggggtcagg caaaaacaga aggagacatt 18480
tcgctaaact ttttcacaac aactgcggca gcagacaata atcccaaagt ggttctttac 18540
agcgaagatg taaaccttca agccccgat actcacttag tatataagcc aacggtggga 18600
gaaaacgtta tcgccgcaga agccctgcta acgcagcagg cgtgtcccaa cagagcaaac 18660
tacatagggt tccgagataa ctttatcggg ttaatgtatt ataacagcac agggaacatg 18720
ggagttctgg caggtcaggc ctgcagtta aacgcagttg tagacctgca agatcgaaac 18780
acggaactgt cctatcagct aatgctagat gctctgggtg acagaactcg atatttctca 18840
atgtggaatc aggccgtgga cagctacgat ccagacgtta ggattatcga gaacctaggg 18900
gtggaagacg agctgccaa ttactgtttt cactcccag gcatgggtat ttttaactcc 18960
tacaaggggg taaaaccaca aaatggcggg aatggtaact gggaagcaaa cggggacctt 19020
tcaaattgcca atgagatcgc tttaggaaac atttttgcca tggaaattaa cctccacgca 19080
aacctgtggc gcagcttctt gtacagcaat gtggcgctgt acctgccaga cagctataaa 19140
ttcactcccg ctaacatcac tctgcccgc aacaaaaaca cctacgagta tatcaacggg 19200
cgcgtcactt ctccaaccct ggtggacacc tttgttaaca ttggagcccg atggtcgcg 19260
gatcccatgg acaacgtcaa cccctttaac catcaccgga acgcgggcct ccgttaccgc 19320
tccatgctgc tgggaaatgg acgcgtgggt cctttccaca tacaagtgcc gcaaaaattt 19380
ttcgcgatta agaacctcct gcttttgccc ggctcctaca cttacgagtg gagcttcaga 19440
aaagacgtga acatgattct gcagagcacc ctgggcaatg atcttcgagt ggacggggcc 19500
agcgtccgca ttgacagcgt caactgtac gccaaacttt tccccatggc gcacaacacc 19560
gcttctacct tggaagccat gctgcgaaac gacaccaacg accagtcgtt taacgactac 19620

ctccagcgcg ccaacatgct ttatcccatt cgggccaacg ccaccaacgt tcccatttcc 19680
 attccctccc gcaactgggc ggccttccgg ggatggagct tcacccgcct taaagccaag 19740
 gaaacgcctt ccttgggctc cggctttgac ccctactttg tgtactcagg caccattcct 19800
 tacctggacg gcagctttta cctcaaccac actttcaaac gtctgtccat catgttcgat 19860
 tcttccgtaa gttggccggg caacgaccgc ctcttgacgc cgaacgagtt cgaaattaag 19920
 cgcattgtgg acggggaagg ctacaacgtg gctcaaagta acatgaccaa agactggttt 19980
 ttaattcaaa tgctcagcca ctacaacatc ggctaccaag gcttctatgt tcccaggggc 20040
 tacaaggatc ggatgtattc tttcttccga aactttcagc ccatgagccg ccagggtgccg 20100
 gatcccaccg ctgccggcta tcaagccgtt cccctgccc aacaacacaa caactcgggc 20160
 tttgtggggg acatggggcc gaccatgcgc gaaggacagc catacccgcc caactacccc 20220
 tatccctga tcggcgctac cgccgtcccc gccattaccc agaaaaagtt tttgtgagac 20280
 cgcgctcatgt ggcgcatacc tttttccagc aactttatgt caatgggggc cctgaccgac 20340
 ctccgacaga acatgcttta cgctaactcc gcccatgccc tggatatgac ttttgagggtg 20400
 gaccccatga acgagcccac gttgctgtac atgctttttg aggtgttcga cgtgggtcaga 20460
 gtgcaccagc cgcaccgcgg tattatcgag gccgtgtacc tgcgcacccc cttctctgcg 20520
 ggcaatgcca ccacataagc cgctgaacta gctggttttt accccagatc ccatgggctc 20580
 cacggaagac gaactgcggg ccattgtgcg agacctgggc tgcggaccct acttctctggg 20640
 cacctttgac aagcggtttc ccgggttcgt gtctcctcgc aaactcgcgt gcgcgatcgt 20700
 gaataccgcc ggccgagaga ccggaggaga gcattggcta gctctgggct ggaacccccg 20760
 ctcgctccacg tttttcctgt tcgaccctt tggcttttca gaccaacgct tgaagcagat 20820
 ctatgcattt gaatatgagg gtctactcaa gcgaagcgcg ctggcctcct ccgccgatca 20880
 ctgtctaacc ctggtaaaga gactcagac ggttcagggc cctcacagcg ccgcctgtgg 20940
 ccttttttgt tgcattgttt tgacgcctt tgtgaactgg ccggacaccc ccatggaaaa 21000
 caacccacc atggacctc tgactggcgt tcccaactcc atgctccaaa gcccagcgt 21060
 gcagaccacc ctctccaaa accagaaaaa tctgtacgcc tttctgcaca agcactctcc 21120
 ctactttcgc cgccatcggg aacaaataga aaatgcaacc gcgtttaaca aaactctgta 21180
 acgtttaata aatgaacttt ttattgaact ggaaaacggg tttgtgattt ttaaaaatca 21240
 aaggggttga gctggacatc catgtgggag gccggaaggg tgggtgttctt gtactggtac 21300
 ttgggcagcc acttaaactc tggaatcaca aacttgggca gcggtatttc tgggaagttg 21360
 tcgtgccaca gctggcgggt cagctgaagt gcctgcagaa catcgggggc ggagatcttg 21420
 aagtcgcagt ttatctggtt cacggcacgc gcgttgcggt acatgggatt ggcacactga 21480
 aacaccagca ggctgggatt cttgatgcta gccagggcc aaggcgtcgt cacgtcaccg 21540
 gtgtcttcta tgttgacag cgaaaaaggc gtgactttgc aaagctggcg tcccgcgcga 21600
 ggcacgcaat ctcccaggta gttgactca cagcggatgg gcagaagaag atgcttgtgg 21660
 ccgcgggtca tgtagggata ggccgctgcc ataaaagctt cgatctgcct gaaagcctgc 21720

ttggccttgt	gcccttcggt	ataaaaaaca	ccgcaggact	tggttgaaaa	ggtattactg	21780
gcgcaagcgg	catcgtgaaa	gcaagcgcgt	gcgtcttcgt	ttcgttaactg	caccacgctg	21840
cggccccacc	ggttctgaat	caccttggcc	ctgccgggggt	tttccttgag	agcgcgctgg	21900
ccggcttcgc	tgccacatc	catttccacg	acatgctcct	tgtaaatcat	ggccagaccg	21960
tggaggcagc	gcagctcctc	gtcatcgtcg	gtgcagtgat	gctcccacac	gacgcagcca	22020
gtgggctccc	acttgggctt	ggaggcctcg	gcaatgccag	aatacaggag	aacgtagtgg	22080
tgcagaaaac	gtcccatcat	ggtgccaaag	gttttctggc	tgctgaaggt	catcgggcag	22140
tacctccagt	cctcgttaag	ccaagtgttg	cagatcttcc	tgaagaccgt	gtactgatcg	22200
ggcataaagt	ggaactcatt	gcgctcggtc	ttgtcgatct	tatacttttc	catcagacta	22260
tgcataatct	ccatgccctt	ttcccaggcg	caaacaatct	tggtgctaca	cgggttaggt	22320
atggccaaag	tggttggcct	ctgaggcggc	gcttgttctt	cctcttgagc	cctctcccga	22380
ctgacggggg	ttgaaagagg	gtgccccttg	gggaacggct	tgaacacggg	ctggcccag	22440
gcgtcccga	gaatctgcat	cgggggattg	ctggccgtca	tggcgatgat	ctgaccccgg	22500
ggctcctcca	cttcgtcctc	ctcgggactt	tcctcgtgct	tttcggggga	cggtagcggga	22560
gtagggggaa	gagcgcggcg	cgcttcttct	ttgggcggca	gttccggagc	ctgctcttga	22620
cgactggcca	ttgtcttctc	ctaggcaaga	aaaacaagat	ggaagactct	ttctcctcct	22680
cctcgtcaac	gtcagaaagc	gagtcttcca	ccttaagcgc	cgagaactcc	cagcgcatag	22740
aatccgatgt	gggctacgag	actccccccg	cgaacttttc	gccgcccccc	ataaactacta	22800
acgggtggac	ggactacctg	gccctaggag	acgtactgct	gaagcacatc	aggcggcgaga	22860
gcgttatcgt	gcaagatgct	ctcaccgagc	gactcgcggg	tccgctggaa	gtggcggaac	22920
ttagcgccgc	ctacgagcga	accctcttct	ccccaaagac	tccccccaag	aggcaggcta	22980
acggcacctg	cgagcctaac	cctcgactca	acttctaccc	tgcttttgcc	gtgccagagg	23040
tactggctac	gtaccacatt	tttttccaaa	accacaaaat	ccctctctcg	tgccgcgcca	23100
accgcaccaa	agccgatcgc	gtgctgcgac	tggaggaagg	ggctcgcata	cctgagattg	23160
cgtgtctgga	ggaagtccca	aaaatctttg	aagggtctggg	ccgcgacgaa	aagcgagcag	23220
caaacgctct	ggaagagaac	gcagagagtc	acaacagcgc	cttggtagaa	ctcgagggcg	23280
acaacgccag	actggccgtc	ctcaaacggg	ccatagaagt	cacgcacttc	gcctaccccg	23340
ccgttaacct	ccctccaaaa	gttatgacag	cgggtcatgga	ctcgtgctc	ataaagcgcg	23400
ctcagccctt	agaccagag	cacgaaaaca	acagtgacga	aggaaaaccg	gtggtttctg	23460
atgaggagtt	gagcaagtgg	ctgtcctcca	acgaccccg	cacgttggag	gaacgaagaa	23520
aaaccatgat	ggccgtggtg	ctagtaccg	tgcaattaga	atgtctgcag	aggttctttt	23580
cccaccaga	gaccctgaga	aaagtggagg	aaacgctgca	ctacacattt	aggcacggct	23640
acgtgaagca	agcctgcaag	atttccaacg	tagaacttag	caacctcatc	tcctacctgg	23700
ggatcttgca	cgaaaaccgc	ctcggacaaa	acgtgctgca	cagcacactg	aaaggagaag	23760

cccgccgaga ctatgtgcga gactgcgtgt tcctagcgct agtgtacacc tggcagagcg 23820
 gaatgggagt ctggcagcag tgcctggagg acgaaaacct caaagagctt gaaaagctgc 23880
 tgggtgcgctc cagaagggca ctgtggacca gttttgacga gcgcaccgcc gcgcgagacc 23940
 tagctgatata tttttttcct cccaagctgg tgcagactct ccgggaagga ctgccagatt 24000
 ttatgagtca aagcatcttg caaaacttcc gctctttcat cttggaacgc tcgggaatct 24060
 tgcccgccac tagctgcgcc ctaccacag attttgtgcc tctccactac cgcgaaatgcc 24120
 caccgccgct gtggccgtac acttacttgc ttaaactggc caactttcta atgttccact 24180
 ctgacctggc agaagacgtt agcggcgagg ggctgctaga atgccactgc cgctgcaacc 24240
 tgtgcacccc ccaccgctct ctagtatgca acactcccct gctcaatgag acccagatca 24300
 tcggtacctt tgaaatccag ggaccctccg acgcggaaaa cggcaagcag ggggtctgggc 24360
 taaaactcac agccggactg tggacctccg cctacttgcg caaatttgta ccagaagact 24420
 atcacgccc ccaaattaaa ttttacgaaa accaatcaaa accacccaaa agcgagttaa 24480
 cggcttgctg cattacgcag agcagcatag ttgggcagtt gcaagccatt aacaaagcgc 24540
 ggcaagagtt tctcctaaaa aaaggaaaag ggggtctactt ggacccccag accggcgagg 24600
 aactcaacgg accctcctca gtcgcaggtt gtgtgcccc tgcgcgccc aaagaacacc 24660
 tcgcagtggg acatgccaga gacggaggaa gaggagtggg gcagtgtgag caacagcgaa 24720
 acggaggaag agccgtggcc cgaggggtgc aacggggaag aggacacgga gggacggcga 24780
 agtcttcgcc gaagaactct cgccgctgcc cccgaagtcc cagccggccg cctcggccca 24840
 agatcccgc caccaccgta gatgggatat caagaccaa aagccgggta agagaaacgc 24900
 tcgccccgc cagggctacc gctcgtggag aaagcacaaa aactgcatct tatcgtgctt 24960
 gctccagtgc ggcggagacg tttcgttcac ccgtagatac ttgcttttta acaaaggggt 25020
 ggccgtcccc cgtaacgtcc tccactacta ccgtcactct tacagctccg aagcggacgg 25080
 ctaagaaaac gcagcagttg ccggcgggag gactgcgtct cagcgcgccga gaacccccag 25140
 ccaccagggg gctccgaaac cgcatatttc ccaccctcta cgctatcttt cagcaaagcc 25200
 gggggcagca gcaagaactg aaaataaaaa accgcacgct gaggtcgctt acccgaagct 25260
 gcctctatca caagagcgaa gagcagctgc agcgaaccct ggaggacgca gaagcgctgt 25320
 tccagaagta ctgcgcgacc accctaaata actaaaaaag cccgcgcgcg ggacttcaaa 25380
 ccgtctgacg tcaccagccg cgcgccaaaa tgagcaaaga gattcccacg cttacatgt 25440
 ggagttacca gccgcagatg ggattagccg ccggcgccgc ccaggattac tccacgaaaa 25500
 tgaactggct cagcgcgggg ccccatatga tttccgcgct aaacgacatt cgcgcccacc 25560
 gcaatcagct attgttagaa caggctgctc tgaccgccac gcccgtaat aacctgaacc 25620
 ctcccagctg gccagctgcc ctggtgtacc aggaacgcc tccaccacc agcgtacttt 25680
 tgccccgtga cgcgcaggcg gaagtccaga tgactaacgc gggcgcgcaa ttagcgggcg 25740
 gatcccggtt tcggtacaga gttcacggcg ccgcacccta tagcccaggt ataaagaggc 25800
 tgatcattcg aggcagaggt gtccagctca acgacgagac agtgagctct tcgcttggtc 25860

tacgaccaga cggagtgttc cagctcgcgg gctcggggccg ctcttcgttc acgcctcgcc 25920
aggcatacct gactctgcag agctctgcct ctcagcctcg ctcgggagga atcggacccc 25980
ttcagtttgt ggaggagttt gtgccctcgg tctactttca gcctttctcc ggatcgcccc 26040
gccagtaccc ggacgagttc atccccaact tcgacgcggt gagtgactct gtggacgggt 26100
atgactgatg tcgagcccgc ttcagtgtta gtggaacaag cgcggctcaa tcacctgggt 26160
cgttgccgcc gccgctgctg cgtggctcgc gacttgagct tagctctcaa gtttgtaaaa 26220
aaccgcgtccg aaaccgggag cgctgtgcac ggggttgagc tagtggggtcc tgagaaggcc 26280
accatccacg ttctcagaaa ctttgtggaa aaaccattt tggttaaacg agatcagggg 26340
ccttttgtaa tcagcttact ctgcacctgt aacctgttg accttcacga ctattttatg 26400
gatcatttgt gcgctgaatt caataagtaa agcgaattct taccaagatt atgatgtcca 26460
tgactgttcc tcgccactat acgatgttgt gccagtaaac tctcttgctg acatctatct 26520
gaactgttcc ttttggtccg cacagcttac ttggtactac ggtgacaccg tcctttctgg 26580
ctcactgggc agctcacacg gaataacact tcacctcttt tcgccgtttc gatacggaaa 26640
ctacagctgt cgtgccggtta cctgcctcca cgttttcaat cttcagccct gtccaccgac 26700
caaacttgta tttgtcgact ctaagcactt acagctcaac tgcagcattc taggccccag 26760
tatcttggtg acatacaata aaatcagggt ggtggaattt gtctactacc caccagcgc 26820
ccgcggtttt ggggaaattc ctttcagat ctactacaac tatcttgcca cacattatgc 26880
aagtcaacag caactaaact tgcaagcacc cttcacgcca ggagagtact cctgtcacgt 26940
aggctcctgc acagaaactt ttattctctt caacagatct tctgccattg aacgcttcac 27000
tactaactac tttagaaacc aagttgtgct tttcactgac gaaaccctta acgtcacct 27060
ggactgtgca tgtttttctc atgacaccgt aacttggaact cttacaata ctctctggct 27120
cgcgttcgat aaccaaagct tgattgttaa aaattttgat ttaaccttta ctaaaccctc 27180
tcctcgcgaa atagttatct ttgctccttt taatccaaaa actaccttag cctgtcaggt 27240
tttgtttaag ctttgccaaa caaactttta gtttgtttat ttgcctccgc aatctgtcaa 27300
actcatagaa aaatacaaca aagcgcctgt cttggctcct aaaaccttct accactggct 27360
aacctacacg gggctgtttg cactaattgt ttttttcta attaacattt ttatatgttt 27420
cttgcccttc tccttctttt cgcgaacacc gttgccgcag aaagacctct cttattact 27480
gtagcgcttg ctatacaaaa ccaagagtgg tcaaccgtgc tctcaatcta ttttcaattt 27540
ttcattttgt ccttaatact ttctcttatt gtcgttaaca atgatctgga gcattggctc 27600
cgcctttttt tggctgctta gtgcaaaagc cactattttt cacaggatatg tggaagaagg 27660
aactagcacc ctctttacga tacctgaaac aattaaggcg gctgatgaag tttcttggtta 27720
caaaggctcg ctctcagacg gcaaccactc attctcagga cagacccttt gcatccaaga 27780
aacttatttt aaatcagaac tacaatacag ctgcataaaa aactttttcc atctctacaa 27840
catctcaaaa ccctatgagg gtatttacia tgccaagggt tcagacaact ccagcacacg 27900

gaactttttac ttttaattctga cagttatttaa agcaattttcc attcctatct gtgagtttag 27960
ctcccagttt ctttctgaaa cctactgttt aattactata aactgcacta aaaatcgcct 28020
tcacaccacc ataattctaca atcacacaca atcaccttgg gttttaaacc taaaattttc 28080
tccacacatg ccttcgcaat ttctcacgca agttaccgtc tctaacataa gcaagcagtt 28140
tggtttttac tatcctttcc acgaactgtg cgaaataatt gaagccgaat atgaaccaga 28200
ctactttact tacattgccca ttggtgtaat cgttgtttgc ctttgctttg ttattggggg 28260
gtgtgtttat ttgtacattc agagaaaaat attgctctcg ctgtgctcct gcggttataa 28320
agcagaagaa agaattaaaa tctctacact ttattaatgt tttccagaaa tggcaaaact 28380
aacgctccta cttttgcttc tcacgccggt gacgcttttt accatcactt tttctgccgc 28440
cgccacactc gaacctcaat gtttgccacc ggttgaagtc tactttgtct acgtgttgct 28500
gtgctgcgtt agcgtttgca gtataacatg ttttaccttt gtttttcttc agtgcatgga 28560
ctacttctgg gtcagactct actaccgcag acacgcgcct cagtatcaaa atcaacaaat 28620
tgccagacta ctcggtctgc catgattgtc ttgtatttta ccctgatttt ttttcacctt 28680
acttgcgctt gtgattttca cttcactcaa ttttggaana cgcaatgctt cgacccgcgc 28740
ctctccaacg actggatgat ggctcttgca attgccacgc ttggggcggt tggacttttt 28800
agtgggtttg ctttgcatga caaatttaag actccatgga cacatggctt tctttcagat 28860
tttccagtta cacctactcc gccgcctccc ccggccatcg acgtgcctca gggtccctca 28920
ccttctccat ctgtctgcag ctactttcat ctgtaatggc cgacctagaa tttgacggag 28980
tgcaattctga gcaaagggct atacacttcc aacgccagtc ggaccgcgaa cgcaaaaaca 29040
gagagctgca aaccatacaa aacaccacc aatgtaaagc cgggatattt tgtattgtaa 29100
aacaagctaa gctccactac gagcttctat ctggcaacga ccacgagctc caatacgtgg 29160
tcgatcagca gcgtcaaacc tgtgtattct taattggagt ttccccatt aaagttactc 29220
aaaccaaggg tgaaaccaag ggaaccataa ggtgctcatg tcacctgtca gaatgccttt 29280
acactctagt taaaacccta tgtggcttac atgattctat cccctttaat taaataaact 29340
tactttaaat ctgcaatcac ttcttcgtcc ttgtttttgt cgccatccag cagcaccacc 29400
ttcccctctt cccaactttc atagcatatt ttccgaaaag aggcgactt tcgccacacc 29460
ttaaagggaa cgtttacttc gctttcaagc tctccacga ttttcattgc agatatgaaa 29520
cgcgccaaag tggaagaagg atttaacccc gtttatccct atggatattc tactccgact 29580
gacgtggctc ctccctttgt agcctctgac ggtcttcaag aaaaccacc tggggctctg 29640
tccctaaaaa tatccaaacc tttaactttt aatgcctcca aggctctaag cctggctatt 29700
ggtccaggat taaaaattca agatggtaaa ctagtggggg agggacaagc aattcttgca 29760
aacctgccgc ttcaaatcac caacaacaca atttcactac gttttgggaa cacacttgcc 29820
ttgaatgaca ataatgaact ccaaaccaca ctaaaatctt catcgcccct taaaatcaca 29880
gaccagactc tgtcccttaa cataggggac agccttgcaa ttaaagatga caaactagaa 29940
agcgtcttcc aagcgaccct cccactctcc attagcaaca acaccatcag cctcaacgtg 30000

ggcaccggac tcaccataaa tggaaacggt ttacaagctg ttcccttaaa tgctctaagt 30060
 cccctaacta tttccaacaa taacatcagc ctgcgctatg gcagttccct gacggtgctt 30120
 aacaatgaac tgcaaagcaa cctcacagtt cactccccct taaaactcaa ctccaacaac 30180
 tcaattttct tcaacactct atctccgttt agaatcgaga atgggtttctt cacgctctat 30240
 ttgggaacaa aatctggctt gctagttcaa aacagtggct taaaagttca agcgggctac 30300
 ggcttgaag taacagacac caatgctctc acattaagat atctcgctcc actgaccatt 30360
 ccagactcgg gctcagaaca aggcatttctt aaagtaaaca ctggacaggg cctaagtgtg 30420
 aaccaagctg gagcgcttga aacatcccta ggaggtggat taaaatatgc tgataacaaa 30480
 ataacctttg atacaggaaa cggactgaca ttatctgaaa ataaacttgc agtagctgca 30540
 ggtagtggct taacttttag agatgggtgcc ttggtagcca cgggaaccgc atttacgcaa 30600
 aactgttga ctacggctga tccgtctccc aactgcacaa ttatacagga ccgcgacaca 30660
 aaatttactt tggcgcttac cattagtggg agccaagtgc tggggacggt ttccattatt 30720
 ggagtaaaag gccccctttc aagtagcata ccgtcagcta ccgttacagt acaacttaac 30780
 ttgattcca acggagccct attgagctcc tcttcaacta aaggttactg ggggtatcgc 30840
 caaggtccct caattgaccc ttaccccata attaatgcct taaactttat gccaaactca 30900
 ctggcttatc ccccgggaca agaaatccaa gcaaaatgta acatgtacgt ttctactttt 30960
 ttacgaggaa atccacaaag accaatagtt ttaaacaatca cttttaataa tcaaaccagc 31020
 gggttttcca ttagatttac atggacaaat ttaaccacag gagaagcatt tgcaatgccc 31080
 ccatgcactt tttctacat tgctgaacaa caataaacta tgtaaccctc accgttaacc 31140
 cgctccgcc cttccatttt attttataaa ccaccgcatc caccttttca gcagtaaaca 31200
 attgcatgtc agtaggggca gtaaaacttt tgggagttaa aatccacaca gggtcttcac 31260
 aagctaagcg aaaatcagtt acacttataa aaccatcgct aacatcggac aaagacaagc 31320
 atgagtcaa agcttccggt tctggatcag atttttgttc attaacagcg ggagaaacag 31380
 cttctggagg attttccatc tccatctcct tcatcagttc caccatgtcc accgtggtca 31440
 tctgggacga gaacgacagt tgtcatcac ctcataagtc accggtcgat gacgaacgta 31500
 cagatctcga agaattgtcct gtcgccgcct ttcggcagca ctgggccgaa ggcgaaagcg 31560
 cccatgttta acaatggcca gcaccgccc cttcatcagg cgcctagtct ttttagcgca 31620
 acagcgcag cgagctcgc taagactggc gcaagaaaca cagcacagaa ccaccagatt 31680
 gttcatgatc ccataagcgt gctgacacca gccatacta acaaattgtt tcactattct 31740
 agcatgaatg tcatatctga tgttcaagta aattaaatgg cggccccctt tgtaaact 31800
 tcccacgtac aacacctcct ttggcatctg ataattaacc acctccgat accaaatata 31860
 tctctgatta atagtgcgcc cgtacactac ccgattaaac caagttgcca acataatccc 31920
 ccctgccata cactgcaaag aacctggacg gctacaatga cagtgcgaaag tccacacctc 31980
 gttgccatgg ataactgagg aacgccttaa gtcaatagtg gcacaactaa tacaacatg 32040

taaatagtgt	ttcaacaagt	gccactcgta	tgaggtgagt	atcatgtccc	agggaaacggg	32100
ccactccata	aacactgcaa	aaccaacaca	tcctaccatc	ccccgcacgg	cactcacatc	32160
gtgcatggtg	ttcatatcac	agtccggaag	ctgaggacaa	ggaaaagtct	cgggagcatt	32220
ttcatagggc	ggtagtgggt	actccttgta	gggggttcagt	cggcaccggg	atctcctcac	32280
cttctgggcc	ataacacaca	agttgagatc	tgatttcaag	gtactttctg	aatgaaaacc	32340
aagtgccttc	ccaacaatgt	atccgatgtc	ttcgggtccc	gcgtcggtag	cgctccttgc	32400
agtacacacg	gaacaaccac	tcacgcaggc	ccagaagaca	gttttccgcg	gacggtgaca	32460
agttaatccc	cctcagtctc	agagccaata	tagtttcttc	cacagtagca	taggccaaac	32520
ccaaccagga	aacacaagct	ggcacgtccc	gttcaacggg	aggacaagga	agcagaggca	32580
gaggcatagg	caaagcaaca	gaatttttat	tccaactggg	cacgtagcac	ttcaaacacc	32640
aggtcacgta	aatggcagcg	atcttggggt	tcctgatgga	acataacagc	aagatcaaac	32700
atgagacgat	tctcaagggt	attaaccaca	gctggaatta	aatcctccac	gcgcacattt	32760
agaaacacca	gcaatacaaa	agcccgggtt	tctccgggat	ctatcatagc	agcacagtca	32820
tcaattagtc	ccaagtaatt	ttcccgtttc	caatctgtta	taatttgcag	aataatgccc	32880
tgtaaatcca	agccggccat	ggcgaaaagc	tcagataatg	cactttccac	gtgcattcgt	32940
aaacacaccc	tcactttgtc	aatccaaaaa	gtcttcttct	tgagaaacct	gtagtaaatt	33000
aagaatcgcc	aggttaggct	cgatgcctac	atcccggagc	ttcatttctca	gcatgcactg	33060
caaatgatcc	agcagatcag	aacagcaatt	agcagccagc	tcatccccgg	tttccagttc	33120
cggagtcccc	acggcaatta	tcaactgaaa	cgtgggacaa	atcgaaataa	catgagctcc	33180
cacgtgagca	aaagccgtag	ggccagtgca	ataatcacag	aaccagcgga	aaaaagattg	33240
cagctcatgt	ttcaaaaagc	tctgcagatc	aaaattcagc	tcatgcaaata	aacacagtaa	33300
agtttgcggt	atagtaaccg	aaaaccacac	gggtcgacgt	tcaaacatct	cggcttacct	33360
aaaaaagaag	cacattttta	aaccacagtc	gcttcttgaa	caggaggaaa	tatggtgcgg	33420
cgtaaaacca	gacgcgccac	cggatctccg	gcagagccct	gataatacag	ccagctgtgg	33480
ttaaacagca	aaacctttta	ttcggcaacg	gttgagggtct	ccacataatc	agcgcgccaca	33540
aaaatcccat	ctcgaacttg	ctcgcgtagg	gagctaaaat	ggccagtata	gccccatggc	33600
acccgaacgc	taatctgcaa	gtatatgaga	gccaccccat	tcggcgggat	cacaaaatca	33660
gtcggagaaa	acaacgtata	caccccgga	tgcaaaagct	gttcaggcaa	acgcccctgc	33720
ggctccctctc	ggtacaccag	caaagcctcg	ggtaaagcag	ccatgccaaag	cgcttaccgt	33780
gccaagagcg	actcagacga	aaaagtgtac	tgaggcgctc	agagcagcgg	ctatatactc	33840
tacctgtgac	gtcaagaacc	gaaagtcaaa	agttcacccg	gcgcgcccga	aaaaacccgc	33900
gaaaatccac	ccaaaaagcc	cgcgaaaaac	acttccgtat	aaaatttccg	ggttaccggc	33960
gcgtcaccgc	cgcgcgacac	gcccgcgccg	ccccgcgctc	ctccccgaaa	cccgcgcgcg	34020
ccacttccgc	gttcccaaga	caaaggtcgc	gtaactccgc	ccacctcatt	tgcatgttaa	34080
ctcggtcgcc	atcttgcggt	gttatattga	tgatg			34115

<210> 7
 <211> 44
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5L

<400> 7
 gcgcacgcgt ctctatcgat gaattccatt ggtgatggac atgc 44

<210> 8
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5ITR

<400> 8
 gcgcatttaa atcatcatca ataataacc tcaaac 36

<210> 9
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5XTOP

<400> 9
 gatacctagg aacgaggagg atttgatatt g 31

<210> 10
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5XBOT

<400> 10
 atgtacgcct ccgcgctcac 20

<210> 11
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5E4

<400> 11
 gatcgaattc ccactctgta ccccatctct g 31

<210> 12
 <211> 31967
 <212> DNA
 <213> Simian adenovirus

<220>

<221> CDS
<222> (13796) .. (15322)

<220>
<221> CDS
<222> (18257) .. (21010)

<220>
<221> CDS
<222> (27192) .. (29015)

```

<400> 12
catcatcata atatacctta tttggaacg gtgccaatat gataatgagg aggcgggggtt      60
aggggtggag tgaggggtggg gtgcgggatga cgcgggcgcg gggcgggggtg ggagtctgac    120
gtggggcgcg ggggtggagcg cgaggggtgag ggcggggcgga gggcggcggg cgcggcggaa    180
ttgacgtaca cggtagtaag tttgagcgga aattaagtga attgggcgtg ttttttgtaa    240
ctttttgacg tacacggtag taagtttgag cggaaattaa gtgaattggg cgtgtttttt    300
gtaacttttt gacgtacacg gtagtaagtt tgagcggaaa ttaagtgaat tgggcgtggt    360
ttttgtaact ttttgacgta cacggtagta agtttgagcg gaaattaagt gaattgggcg    420
tgttttttgt aactttttga cgtacacggg agtaagtttg agcggaaatt aagtgaattg    480
ggcgtgtttt ttgtaacttt ttgacgtaca cggtagtaag tttgagcgga aattaagtga    540
attgggcgtg ttttttgtaa ctttttgacg tacacggtag taagtttgag cggaaattaa    600
gtgaattggg cgtgtttttt gtaacttttt ggtcattttg gcgcgaaaac tgagtaatga    660
ggaagtgaga cggactctgc ctttttttac ggttgggagg gaaaactgct gatcagcgct    720
gaactttggg ctctgacgcg gtggtttccc tacgtggcag tgccacgaga aggctcaaag    780
tcctcgtttt attgtgtgct cagccttttt gaggggtattt aaacaccgtc agaccgtcaa    840
gaggccactc ttgagtgcga gcgagtagag ttttctcctc cgtcgctgcc gcggctgctc    900
agtcttaccg ccaggatgcg aatgctgccg gagatcttca ccgggtcctg ggaagatggt    960
ttccaggggac ttttagaatc tgaagacaac tttccccaac ctctgagcc ggaggagcta   1020
cctgaggttt cgcttcacga tctgtttgac gtggagggtg agagccccga cggagatccg   1080
aacgaggaag ctgttgatgg tatgttcccc gactggatga tatctcagag cgagagtgct   1140
gaaggcagtg cggactcggg cgtttctggg gttggaaacc tgggtggagggt ggatctggac   1200
ttgaagtgtt acgaggaagg ttttcctcct agcgactcag agactgatga agcctcagaa   1260
gcggaaggtc aagaggagtc tgtgtgtggt tatgtgaaga ttaatgaggg ggagaacctg   1320
ctggtgttgg actgtccgga ccaacctgga catggctgtc gagcctgtga ctttcaccgg   1380
gggaccagcg gaaacccgga agctatctgt gctttgtgct acatgcgtct gaacgagcac   1440
tgcataataca gtgagtgtta ttcattgggtt atttatgggg aaagttgggg gaaagtcttg   1500
agaaggggaa aagtttaaca tgtcattttt gtacttgata ggtccagttt cggacgctga   1560
gggggattct gagtcccctg ctggctcctt ccagccctca ccctgctctt tgaccgccac   1620
gcccgcacct gacctagtta gaccaacgcc ctgccgagtg tcctgtagac gacgtgcagc   1680
tgtaattgac atagaagatt tattggcccc tgatgacgag aacgcacctt tgaacctgtg   1740

```


cctgaaacgc	cctaagacat	cttgagtgtt	tatgctgtta	ataaaagtgt	tgacccttag	1800
atcctgtgtt	tattccttgg	gcgtgtgcgc	gggtatataa	agcagctgcg	ggctggagtg	1860
ttagtttatt	ctgatggagt	actggagtga	gctgcagaat	taccagagcc	tccggcgcct	1920
gctggagttg	gcctctgcca	gaacatccac	ctgctggagg	ttctgttttg	gctcgactct	1980
cagtaacgtg	gtgtatcggg	tgaagcaaga	gtacagctcg	cgcttttctg	agctgttggc	2040
ccgctacccg	gctgtttttg	tttctctgga	tctaggccat	cacgtttatt	tccaagaagc	2100
tgtagtcaga	tatttggaatt	tttctactcc	cgggcgctgcg	gtttctgcga	ttgccttcat	2160
ctgctttgtg	ctagatcgat	ggagcgccca	aaccgcctg	agcccggggt	acaccctgga	2220
ctacctgacc	atgtccctgt	ggagggccat	gctgcggaag	aggaggggtct	caggcttctc	2280
gccggcgcg	cctccgcacg	gactggatcc	ggtgctggag	gagtcggagc	tggaggagga	2340
ggagaacccg	agggccggcc	tggaccctcc	ggcggaatag	tgacggaacc	ggaggatccc	2400
caagagggtg	ctagtcaggg	gggagggggg	ccgaagagaa	agcgggatga	agaggaggcg	2460
atggaccccc	acaggtttct	aaaagaactg	actttaagct	taatgtctaa	gagaagaccc	2520
gagacggtgt	ggtggtctga	tttgagagaag	gagttccacc	agggggagat	gaatctgttg	2580
tacaagtatg	ggtttgagca	ggtgaagact	cactggctgg	aagcctggga	ggactgggag	2640
atggctttta	acatgtttgc	caaggtggcg	ctgcgcccgg	acactattta	caccgtgact	2700
aagacggtgg	aaatccgcaa	gcctgtgtat	gtgattggca	acggggccgt	ggttcggttc	2760
cagaccaccg	accgggtggc	ctttaattgc	tgtatgcaga	acctgggccc	gggggtgatt	2820
aatcttaatg	gagtgacctt	ttgcaatgtc	agattcgcg	gggatggatt	caacgggacg	2880
gtgtttgccg	ccaccacca	gataacccta	cacggggtgt	tcttcagca	tgtaggcggg	2940
gcttgtgtag	atacctgggc	gagggcctct	gtgaggggct	gcacctttgt	gggctgttgg	3000
aaagcggtgg	tgggtcgacc	caagagtgtg	ctgtctgtga	agaaatgtgt	gtttgagaga	3060
tgtctgatgg	ccatggtggt	ggagggccag	ggtaggatcc	gccataacgc	gggctccgag	3120
aatacctgtt	ttgccctgct	gaagggtagc	gcgaccgtga	agcataacat	gatctgcggg	3180
gtgggtcact	cgcagctgct	gacctgtgcg	gatggcaact	gccaggccct	gcgcacggtg	3240
catgtggtgt	cccaccggcg	ccgcccctgg	ccggtgtttg	aacataacat	gctgatgcgc	3300
tgtaccatgc	acctgggcta	ccgccgcggc	gtgtttgtgc	cccatcagt	taacctgacc	3360
cacaccaagg	tgttgctgga	gacggatgct	ttttcgcgag	tgaatctgaa	tgggggtgttc	3420
gatctgacta	tggagatgta	caagatagtg	agatttgatg	aatcaaagac	cgttgtcgc	3480
ccctgcgagt	gcggtgccaa	tcacctgagg	atgtatcccg	tgaccctgaa	cgtgacggag	3540
gagctgcgcc	cggaccacca	gatgctgtcc	tgtctgcgca	ccgattacga	aagcagcgat	3600
gaggattaag	aggtgagggg	cggggcttgc	atggggata	aaggtggggg	aggaggtggg	3660
gagggggaaa	acccaaaatg	agcggatcga	tgggaaggag	cgctgtgagt	tttgagggcg	3720
gggtgttcag	cccatatctg	acaaccgctc	tccccgcctg	ggcaggagtg	cgtcagaatg	3780

tggtgggctc	caacgtggac	ggacgtccg	tgccccctgc	caactccgcc	actctcacct	3840
acgccaccgt	cggatcgtcg	ctggacaccg	ccgctgccgc	cgccgcttca	gccgccgctt	3900
ctactgctcg	cggtatggca	gctgatttcg	gactgtatca	gcaactggct	gcgctcgct	3960
cgtcgctgag	agaagatgat	gccctgtccg	tggtgctgac	ccgcctggag	gagctgtccc	4020
agcagctgca	agagctgtct	gccaaagtgg	atgcacagaa	cgccccgct	acccaatgaa	4080
taaataaacg	agacaccgag	tgtgtttgga	aatcaaaatg	tgtttttatt	tgttttttct	4140
ggcgcggtag	gcccttgacc	acctgtcgcg	gtcgttaagg	accttggtga	tgttttccag	4200
cacccggtag	aggtgggctt	ggatgttgag	gtacatgggc	atgagcccgt	ctcgggggtg	4260
gaggtagcac	cactggaggg	cgtcgtgctc	gggggtggtg	ttgtagataa	tccagtcgta	4320
gcagggtttt	tgggcatgga	agcggaagat	gtctttgaga	agcaggctga	tggccagggg	4380
gaggcccttg	gtgtaggtgt	tcacaaagcg	gttgagctgg	gagggatgca	tgcgggggga	4440
gatgagatgc	atcttggcct	gaatcttgag	gttggcgatg	ttgccgcca	gatcccgcg	4500
ggggctcatg	ttgtgcagga	ccaccaggac	ggtgtagccg	gtgcacttgg	ggaatttgtc	4560
atgcaacttg	gaaggggaagg	cgtggaagaa	cttgagagacc	cccttggtggc	cgccgaggtt	4620
ctccatgcat	tcgtccatga	tgatggcgat	gggaccctg	gcggccgccc	tggcgaagac	4680
gttgctcggg	tgggagacgt	cgtagttctg	ttccaggggtg	agctcgtcgt	aggccatttt	4740
gacgaagcgg	gggagcaggg	tgcccactg	ggggacgatg	gtaccttcgg	gaccgggggc	4800
gtagttgccc	tcgcagattt	gcatctccca	ggccttgatc	tccgaggggg	ggatcatgtc	4860
cacctggggc	gcatgaaga	agacgggtctc	cggggcgggg	ttgatgagct	gggaggagag	4920
gaggttgccg	agcagctgcg	acttgccgca	cccgggtggc	ccgtagatga	ccccgatgac	4980
gggttgccagc	tggtagttaa	aggagctgca	gctgccgtcc	tcgcgcagga	acggggcgac	5040
ctcgttcac	atgcttctga	cgtgatgggt	ttccctgacg	aggtcttgca	agagccgctc	5100
gccgcccagg	gagagaagct	cttccaggct	gcggaaatgc	ttgaggggtt	tgaggccgtc	5160
ggccatggtc	atcttttcca	gggactggcg	gagcaggtac	aggcggctcc	agagctcgg	5220
gacgtgttct	acggcatctc	gatccagcag	acttcttgg	tgcggggggt	ggggcggtt	5280
tggctgtagg	ggaccagccg	gtgcgcgtcc	agggaggcga	gggtgacgtc	tttccagggc	5340
cgcagcgttc	gcgtgaggg	ggtctcggtg	acggtgaagg	gatgcgtcc	cggttgggcg	5400
ctggccaggg	tcctcttgag	actcatcctg	ctggtgtgga	agcgggcgtc	ttctccctgg	5460
gagtcggcca	ggtagcattt	gagcatgagg	tcgtagctga	gggcctcggc	cgcgtggccc	5520
ttggcgcgca	gcttgccctt	ggagacgtgt	ccgcaggcgg	gacagtgcag	gcacttgagg	5580
gcgtagagct	tggggggccag	gaagacggac	tcgggggagt	aggcgtcggc	gccgcactga	5640
gcgcacgtgg	tctcgcactc	gacgagccag	gtgagctccg	ggtgttgggg	atcaaaaacc	5700
agctggcccc	cgtgtttttt	gatgcgcttc	ttacctcggg	tctccatgag	gcggcgctcc	5760
gcttcggtga	cgaagaggct	gtcgggtgtc	ccgtagacgg	atttgagcgc	gcgctgctcc	5820
aggggaatcc	cgcgatcctc	cgcgtgcagg	aactcggacc	actctgagac	gaaggcccgg	5880

gtccacgcga	ggacaaagga	ggcgatctgg	gacgggtagc	ggtcgttctc	caccagggga	5940
tccaccttct	ccaggggtgtg	caggcagagg	tcgtcctcct	ccgcgtccat	gaagggtgatt	6000
ggcttgtaag	tgtatgtcac	gtgaccgtcg	gggtcgcgcg	tgggcttata	aaagggggcg	6060
tgcccgccct	ccccgtcact	ttcttccgca	tcgctgtgga	cgagatccag	ctgctcgggt	6120
gagtaggcgc	gctggaaggc	gggcatgacc	tcggcgctga	gggtgtcagt	ttccacgaac	6180
gaggtggatt	tgatattgac	ctgtccggcg	gcgatgcttt	tgacgggtggc	ggggtccatc	6240
tggtcagaaa	agacgatctt	tttgttgctc	agcttgggtg	cgaacgacct	gtagagggcg	6300
ttggagagca	gcttggcgat	ggagcgcagg	gtctggttct	tctcgcggtc	ggcgcgctcc	6360
ttggcggcga	tggttagctg	gacgtactcg	cgggccacgc	agcgccattc	ggggaagacg	6420
gtggcgcgct	cgtccggcag	gaggcgcacg	cgccagccgc	ggttggtgag	ggtgatgagg	6480
tccacgctgg	tggccacctc	gccgcgcagt	ggctcgttgg	tccagcagag	gcgcccgcgc	6540
ttgcgcgagc	agaagggggg	caggacgtcg	agctggctct	ccgcgggggg	gtcggcgctcg	6600
atggtgaaga	tgcccggtag	caggtggcgg	tcgaagtagt	cgatggcgac	cgcggggtcg	6660
gcgagggcgc	gttcccagtc	cctgaccgcc	agggcgcgct	cgtaggggtt	gaggggcgcc	6720
ccccagggca	tgggatgggt	gagggccgag	gcgtacatgc	cgagatgtc	gtagacgtag	6780
aggggctcgc	ggagcacgcc	gaggtaggtg	ggatagcagc	gtccgccgcg	gatgctggcg	6840
cgcacgtagt	cgtacatctc	gtgcgagggg	gcgaggaggc	cgctccgag	gtcgccgcgc	6900
tgcggtctga	cggcccggta	ggtgacctgg	cggaagatgg	cgtgcgagtt	ggaggagatg	6960
gtgggcccgt	ggaagatggt	gaagctggcc	tcggggagtc	cgacggcgct	gtggacgaac	7020
tgggcgtagg	agtcgcgcag	cttctgcacg	agcgcggcgg	tgacgagcac	gtccagggcg	7080
cagtagtcga	gggtctcgcg	gacgaggtcg	taacggggct	cttgcttctt	ttcccagagt	7140
tcgcggttga	ggaggtactc	ctcgcgatcc	ttccagtact	cttcggccgg	aaagccgcgt	7200
tcgtccgcca	ggtaagaacc	cagcatgtag	aagcggttga	cggctcggta	gggacagcag	7260
cccttctcga	cgggcaggga	gtaggcctgc	gcggccttcc	tgagcgaggt	gtgggtgagg	7320
gcgaaggtgt	cgcgccacct	gaccttgagg	aactggaacc	tgaagtcggt	gtcgtcgcag	7380
gcgccccgct	cccagagccc	gtagtcgggt	cgtttctggc	tgcggggggt	gggcagggcg	7440
aagggtgacgt	cgttgaagag	gatcttgccg	gcgcgcggca	tgaagttgag	ggtgatcctg	7500
aagggccccg	gcacgtccga	gcggttggtt	atgacctggg	ccgcgaggac	gatctcgtcg	7560
aagccgttga	tggttgaggc	gacgatgtag	agctcgacga	agcgcggggc	cccctgcagc	7620
ttgggggcct	tcttgagctc	ctcgtaggtg	aggcagtcgg	gcgagtagag	gcccagctcc	7680
tgctggggcc	attcggccac	ctggggggtg	gcttgcaaga	agccccgcca	gagctgcagg	7740
gcgagctggg	tctggaggcg	gtcgcggtag	tcgcggaact	ttttgcccac	cgccatcttc	7800
tcgggggtga	ccacgtagaa	ggtgcggccg	tcctggcccc	aggcgtccca	gttctgctcg	7860
cgggcgagac	ggcaggcctc	ctcgacgagg	gcctcctccc	cggagagatg	catgactagc	7920

atgaagggga	cgagttgctt	gccgaaggca	cccatccacg	tgtagggtctc	tacgtcgtag	7980
gtgacgaaga	gacgttcggt	gcgaggatgc	gagccgagag	gaaagaagtt	gatctcctgc	8040
caccagccgg	aggagtgggc	gttgacgtgg	tggaagtaga	agtcacgccg	gcggaccgtg	8100
cattcgtgct	gatatttgta	aaagcgggcg	cagtactcgc	agcgctgcac	gctctgcact	8160
tcctgaacga	gatgcacccg	gcgcccgcgc	accaggaggc	ggagggggca	gtccagtgga	8220
gcttcggcgc	gctgtccttc	agcctcgtca	tgctcttctg	cacctgcacg	ctcctgctgt	8280
gggtggagga	cggagggagt	gacgacgccg	cgcgagccgc	aggccagat	gtcgacgcgc	8340
ggcggcctga	ggctcagcgc	caggggtgcg	atctgagcgg	cgccaggga	gtcgaggaag	8400
gcctcgtga	ggtcgacggg	cagcgtccgc	cggtggactt	gcaggagacg	ggtaagggcc	8460
ggcgccaggc	gctgatggta	cttgatctcg	agcggttcgt	tggtggagggt	gtcgatggcg	8520
tagagcaggg	cctgaccgcg	ggcggcgcgc	atggtgccgc	ggtgccggcg	gtaggtggcg	8580
tattcggggg	ggctcgttac	atcaccgcgc	tgggcctggc	gccgggcggc	agcgggggtt	8640
ctggtcccgc	cggcatgggc	ggcagcggca	cgtcggcgcg	gggctccggc	agcggctggt	8700
gctgagctcg	cagctgactg	gcgtgcgcga	cgacgcggcg	gttgagggtcc	tggatgtgcc	8760
tccgctgcgt	gaagaccacc	ggtccccgga	ctcggaacct	gaaagagagt	tcgacagaat	8820
caatctcggc	atcgttgacg	gccgcctgac	gcaggatctc	ctgcacgtcg	cccagattgt	8880
cctggtaggc	gatctcggac	atgaactggt	cgatctcttc	ctcctggagt	tcgccgcgtc	8940
cggcgcgttc	gacggtggcc	gcgaggtcgt	tggagatgcg	agccatgagc	tgggagaagg	9000
cgttgaggcc	gttctcgttc	cacacgcgac	tgtagacgac	gttgccgacg	gcgtcccggg	9060
cgcgcatgac	cacctgcgcg	acgttgagct	ccacgtgtcg	cgcaagacg	gcgtagttgc	9120
gcaggcgctg	gaagaggtag	ttgagggtag	tggcgatgtg	ctcgacgacg	aagaagtaca	9180
tgaccagcg	gcgacgcgtc	atctcgttga	tgtctccgag	ggcttccaag	cgctccatgg	9240
cctcgtagaa	gtcgacggcg	aagttgaaga	actgggagtt	gcgcgccgcg	accgtcagct	9300
cgtcttgcaa	gagccggatc	agctgggcca	cggtctcccg	cacctcgcgt	tcgaaggccc	9360
ccggcgcttc	ttctctctct	ggttctctcg	cggcctcttc	ttccatgacg	gcttctctct	9420
cctccggttc	ctcgggcacg	ggcctccggc	ggcgacggcg	cctgatgggc	aggcgggtcca	9480
cgaagcgttc	gatgatctct	ccgcggcggc	ggcgcatggt	ttcggtgacg	gcgcggccgt	9540
tctctcgggg	ccgcagttcg	aagacgcccc	cgcgaggccg	gccggcgccg	ccgagagggg	9600
gcaggaggtg	ggggccttcg	ggcagcgaga	gggcgctgac	gatgcaccgt	atcatctggt	9660
gcgtaggtac	agctctccag	gagtcgttga	gcgagtccag	ttggacggga	tccgagaact	9720
tttcgaggaa	agcttcgatc	caatcgcagt	cgcaaggtaa	gctgaggacg	gtgggatgag	9780
gggcttgggc	ggaggcggag	gcggcagaag	aggaggagga	gggcaggctg	gaggtgatgc	9840
tgctgatgat	gtaattgaag	taggcggttt	tcaaacggcg	gatggtggcg	aggaggacga	9900
cgtctttggg	cccggcctgc	tggatgcgca	ggcggtcggc	catgccccag	gcgtggctct	9960
ggcatcggcg	caggtccttg	tagtagtctt	gcatgagtct	ctcgacgggg	acgtcgtctt	10020

cgctcgccccg gtcggccatg cgggtggagc cgaacccgcg caggggctgc agcagggcca 10080
ggctcggcgac cacgcgttcg gccagcacgg cctgctggat ctgggtgagc gtggtctgga 10140
agtcgtccag gtccacgaag cgggtgtagg agcccgtgtt gatggtgtag gtgcagttgg 10200
ccatgacgga ccagttgacg acttgcatgc cgggctgggt gatctcggtg tagcggaggc 10260
gcgagtaggc ccgcgactcg aagacgtagt cgttgacagg gcgcacgagg tactggtagc 10320
cgacgaggaa gtgcggcggc ggctcgcggg agaggggcca gcgcacgggt gcgggggagc 10380
cggggggccag gtcctccagc atgaggcggg ggtagtggta gacgtagcgc gagagccagg 10440
tgatgccggc ggcggagggt gcggcgcggg cgaagtcgcg gacgcggttc cagatgttgc 10500
gcaagggggc gaagcgctcc atggtgggca cgctctggcc ggtgaggcgg gcgcagtcct 10560
gcacgctcta gacgggacag agagcgggag gtttagcggc ccgctccgtg gcctggggga 10620
cagaccgcca ggggtgcgac gcggggaacc ccggttcgag accggctgga tccgtccgtc 10680
cccgcgcgc cggccccgcg tccacgacct caccagaggc cgagaccag ccgcggtgcc 10740
cggacccag atacggaggg gagccttttt gtggtttttt cccgtagatg catccggtgt 10800
tgcgacagat gcgtccgtcg ccagcgccgc cgacgcagcc gccgctccc cccccacta 10860
gcgccgcgga ggctctgtcc ggcggccgcg gcgacccgga ggaggaggcc atcctcgact 10920
tggaagaagg cgagggcctg gcccggtgg gagcgccctc ccccgagcgc catccccgcg 10980
tgcagctggc gagagactcg cgccaggcct acgtgccgcc gcagaatctg ttcagggacc 11040
gcagcggcca ggagcccag gagatgaggg accgcagggt tcacgcgggg cgggagctgc 11100
gcgcgggctt cgaccgtcgg cgggtgttgc gcgccgaaga cttcgagccc gacgagcgca 11160
gcggagtaag tccggcacgg gcgcacgtgt cggcggccaa cctggtgacc gcgtacgagc 11220
agacggtgaa cgaggagcgg agctttcaga aaagcttcaa caaccacgtg cgcaccctga 11280
tcgcgcgcga ggaggtggcc atcggcctga tgcattctgt ggactttgtg gaggcgtacg 11340
tgcagaaccc gtcgagcaag ccgctgacgg cgcagttgtt cctgatcgtg cagcacagtc 11400
gggacaacga gacgttccgc gaggcgatgc tgaacatcgc ggagcccag ggccgctggc 11460
tcttgacct gattaacatc ctgcagagca tcgtggtgca ggagcgcagc ctgagcctgg 11520
ccgacaaggt ggcggccatc aactacagca tgttgagcct gggcaagttt tacgcccga 11580
agatctacaa gagcccctac gtgcccatag acaaggagg gaagatcgac agcttttaca 11640
tgcggtggc gctgaaagt ctgacgtga gcgacgatct gggggtgtac cgcaacgacc 11700
gcatccacaa ggccgtgagc gccagccgcc ggcgcgagct gagcgaccgc gagctgatgc 11760
acagcctgcg gagggcgctg gcgggcgccg gcggcgccga ggaggccag tcctacttcg 11820
acatgggggc ggacttgag tggcagccca gcgcgcgggc cctggaggcg gcgggctacc 11880
gcggcgccgg cggcgtggtc gaggcggagg acgaggacga ggtggagtac gaggaggagg 11940
actgatcggc gaggtgtttt cgtagatgca gcgcgcgac gcggcgccga gcgggcccga 12000
gggggacccc gccgtgctgg cggccctgca gagccaacct tcgggcgtga acgcctccga 12060

tgactgggcg gcggccatgg accgcattttt ggccttgacc acccgcaacc ccgaggcctt 12120
 tagacagcag ccgcaggcca accgccttttc ggccatcttg gaagccgtgg tgccctcgcg 12180
 caccaacccc acgcacgaga aggtcctggc ggtggtgaac gcgctgctgg agagcaaggc 12240
 gatccgcaag gacgaggcgg ggctgattta caacgccctg ctggagcggg tggcgcgcta 12300
 caacagcacc aacgtgcagg ccaacctgga ccgtctgacg acggacgtgc gggaggcggg 12360
 ggcgcagcgg gagcgcttca tgcgcgacac gaacctgggc tcgcagggtg ccctgaacgc 12420
 cttcctgagc acgcagccgg ccaacgtgcc gcgcgggcag gaggactacg tcagtttcat 12480
 cagcgcgctg cgctccttg tggccgaggt gccgcagagc gaggtgtacc agtcgggtcc 12540
 ggactacttc ttccagacct cgcggcaggg cctgcagacg gtgaacctga cgcaggcctt 12600
 caagaacctg gaaggcatgt ggggcgtgcg gggccccgtg ggcgaccggg cgacgatctc 12660
 cagcttgctg acgccgaaca cgcggtgct gctgctgctg atcgcgccct tcaccaatag 12720
 cagtaccatc agccgcgact cgtacctggg ccacctgac acgctgtacc gcgaggccat 12780
 cgggcaggcg cagggtggacg agcagacctt ccaggagatt acgagcgtga gccggggcct 12840
 ggggcagcag gacacgggta gcctggaggc gacgctgaat tttctgctga ccaaccggcg 12900
 gcagaagatc ccctcccagt acacgctgag cacggaggag gagcgcatct tgcgctacgt 12960
 gcagcagtcg gtgagcctgt atctgatgcg cgagggggcg agccccctcg cggcgctgga 13020
 catgacggcc cgtaacatgg agccgtcgct gtacgcggcc caccggccgt tcgtgaaccg 13080
 cctgatggac tacctgcacc gcgccgccgc catgaacggc gagtacttta cgaacgccat 13140
 cctgaacccg cactggatgc cgccgtccgg tttctacacg ggggactttg acatgcccga 13200
 gggcgacgac gggttcctgt gggacgacgt gtcggacagc gtgttcgcgc cgggtgcgtcc 13260
 gggcaagaag gagggcgcg acgagctgcc gctgtccgtg gtggaggcgg cgtcgcgcgg 13320
 ccagagcccc ttccccagcc tcccgtcggt gtcggcgagc agcagcagcg gccgggtctc 13380
 gcgcccgcgg ctggagggcg actacctgaa cgaccgcgtg ctgcgccccg cccggcccaa 13440
 gaactttccc aacaacgggg tggagagcct agtgataag atgaatcgct ggaagaccta 13500
 cgcccaggag cagcgggagt gggaggagag tcagccccgc cccctgcctc cgccgcgtc 13560
 cagggtggcg cggcggggaa aagaccgga agactcggcg gacgatagca gcgtgttgga 13620
 cttggggggg accggtgccg cctcgacaaa cccgttcgcc cacctgcgcc cgaggggcg 13680
 gctgggtcgg ctgtattgag gaaagaaact aataaaagaa aaaagagctt gcttaccaga 13740
 gccatggtcg cagcgtcggg ccctttgtgt gtgttttctc ctccccggta gcgaa atg 13798
 Met
1
 agg cgc gcg gtg gga gtg ccg ccg gtg atg gcg tac gcc gag ggt cct 13846
 Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly Pro
5 10 15
 cct cct tct tac gaa acg gtg atg ggc gcc gcg gat tcg ccg gcc acg 13894
 Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala Thr
20 25 30
 ctg gag gcg ctc tac gtc cct ccc cgc tac ctg ggg cct acg gag ggg 13942

Leu	Glu	Ala	Leu	Tyr	Val	Pro	Pro	Arg	Tyr	Leu	Gly	Pro	Thr	Glu	Gly	
	35					40					45					
agg	aac	agc	atc	cgt	tac	tca	gag	ctg	gcg	ccg	ctg	tac	gac	acc	acc	13990
Arg	Asn	Ser	Ile	Arg	Tyr	Ser	Glu	Leu	Ala	Pro	Leu	Tyr	Asp	Thr	Thr	
50					55					60					65	
cgc	gtg	tac	ctg	gtg	gat	aac	aag	tcg	gcg	gac	atc	gcg	tcg	ctg	aac	14038
Arg	Val	Tyr	Leu	Val	Asp	Asn	Lys	Ser	Ala	Asp	Ile	Ala	Ser	Leu	Asn	
				70					75					80		
tac	cag	aac	gac	cat	agc	aac	ttt	ctg	acc	acg	gtg	gtg	cag	aac	aat	14086
Tyr	Gln	Asn	Asp	His	Ser	Asn	Phe	Leu	Thr	Thr	Val	Val	Gln	Asn	Asn	
			85					90					95			
gac	ttt	acc	ccg	gtg	gag	gcg	ggc	acg	cag	acc	ata	aat	ttc	gac	gag	14134
Asp	Phe	Thr	Pro	Val	Glu	Ala	Gly	Thr	Gln	Thr	Ile	Asn	Phe	Asp	Glu	
		100					105					110				
cgc	tcg	cgg	tgg	ggc	ggc	gac	ctg	aaa	acc	atc	ctg	cgc	acc	aac	atg	14182
Arg	Ser	Arg	Trp	Gly	Gly	Asp	Leu	Lys	Thr	Ile	Leu	Arg	Thr	Asn	Met	
	115					120					125					
ccc	aac	atc	aac	gag	ttc	atg	tcc	acc	aac	aag	ttc	agg	gcc	cgg	ttg	14230
Pro	Asn	Ile	Asn	Glu	Phe	Met	Ser	Thr	Asn	Lys	Phe	Arg	Ala	Arg	Leu	
130					135					140					145	
atg	gta	gag	aaa	gtg	aac	aag	gaa	acc	aat	gcc	cct	cga	tac	gag	tgg	14278
Met	Val	Glu	Lys	Val	Asn	Lys	Glu	Thr	Asn	Ala	Pro	Arg	Tyr	Glu	Trp	
				150					155					160		
ttt	gag	ttc	acc	ctg	ccc	gag	ggc	aac	tac	tcg	gag	acc	atg	acc	ata	14326
Phe	Glu	Phe	Thr	Leu	Pro	Glu	Gly	Asn	Tyr	Ser	Glu	Thr	Met	Thr	Ile	
			165				170						175			
gac	ctg	atg	aat	aac	gcg	atc	gtg	gac	aac	tac	ttg	gaa	gtg	ggg	cgg	14374
Asp	Leu	Met	Asn	Asn	Ala	Ile	Val	Asp	Asn	Tyr	Leu	Glu	Val	Gly	Arg	
		180					185					190				
cag	aac	ggg	gtg	ctg	gag	agc	gac	atc	ggg	gtg	aag	ttt	gac	acg	cgc	14422
Gln	Asn	Gly	Val	Leu	Glu	Ser	Asp	Ile	Gly	Val	Lys	Phe	Asp	Thr	Arg	
	195					200					205					
aac	ttc	cgg	ctg	ggc	tgg	gac	ccg	gtc	acc	aag	ctg	gtc	atg	ccc	ggc	14470
Asn	Phe	Arg	Leu	Gly	Trp	Asp	Pro	Val	Thr	Lys	Leu	Val	Met	Pro	Gly	
210					215					220					225	
gtg	tac	acc	aac	gag	gcc	ttc	cac	ccc	gac	atc	gtc	ctg	ctg	ccc	ggc	14518
Val	Tyr	Thr	Asn	Glu	Ala	Phe	His	Pro	Asp	Ile	Val	Leu	Leu	Pro	Gly	
				230					235					240		
tgc	ggc	gtg	gac	ttc	acg	cag	agc	cgg	ctg	agc	aac	ctg	ctg	ggg	atc	14566
Cys	Gly	Val	Asp	Phe	Thr	Gln	Ser	Arg	Leu	Ser	Asn	Leu	Leu	Gly	Ile	
			245					250					255			
cgc	aag	cgg	atg	ccc	ttc	cag	gcg	ggt	ttt	cag	atc	atg	tac	gag	gac	14614
Arg	Lys	Arg	Met	Pro	Phe	Gln	Ala	Gly	Phe	Gln	Ile	Met	Tyr	Glu	Asp	
		260					265					270				
ctg	gag	ggc	ggc	aac	atc	ccc	gcc	ttg	cta	gac	gtg	gcg	aaa	tac	gag	14662
Leu	Glu	Gly	Gly	Asn	Ile	Pro	Ala	Leu	Leu	Asp	Val	Ala	Lys	Tyr	Glu	
	275					280					285					
gcc	agc	att	cag	aag	gcg	cgg	gag	cag	ggc	cag	gag	atc	cgc	ggc	gac	14710
Ala	Ser	Ile	Gln	Lys	Ala	Arg	Glu	Gln	Gly	Gln	Glu	Ile	Arg	Gly	Asp	
290					295				300					305		
aac	ttt	acc	gtc	atc	ccc	cgg	gac	gtg	gag	atc	gtg	ccc	gtg	gag	aag	14758
Asn	Phe	Thr	Val	Ile	Pro	Arg	Asp	Val	Glu	Ile	Val	Pro	Val	Glu	Lys	

310								315						320						
gat Asp	agc Ser	aag Lys	gac Asp 325	cgc Arg	agt Ser	tac Tyr	aac Asn	cta Leu 330	ctc Leu	ccc Pro	ggc Gly	gac Asp	cag Gln 335	acc Thr	aac Asn	14806				
acg Thr	gcc Ala	tac Tyr 340	cgc Arg	agc Ser	tgg Trp	ttc Phe	ctg Leu 345	gcc Ala	tac Tyr	aac Asn	tac Tyr	ggc Gly 350	gac Asp	ccc Pro	gag Glu	14854				
aag Lys	ggc Gly 355	gtc Val	agg Arg	tcc Ser	tgg Trp	acg Thr 360	ctg Leu	ctg Leu	acc Thr	acc Thr	acg Thr 365	gac Asp	gtc Val	acc Thr	tgc Cys	14902				
ggc Gly 370	tcg Ser	cag Gln	cag Gln	gtg Val	tac Tyr 375	tgg Trp	tcg Ser	ctc Leu	ccg Pro	gac Asp 380	atg Met	atg Met	caa Gln	gac Asp	ccc Pro 385	14950				
gtg Val	acc Thr	ttc Phe	cgg Arg	ccc Pro 390	tcc Ser	agc Ser	caa Gln	gtc Val	agc Ser 395	aac Asn	tac Tyr	ccc Pro	gtg Val	gtg Val 400	gga Gly	14998				
gtc Val	gag Glu	ctc Leu	ctg Leu 405	ccg Pro	gtg Val	cac His	gcc Ala	aag Lys 410	agc Ser	ttt Phe	tac Tyr	aac Asn	gag Glu 415	cag Gln	gcc Ala	15046				
gtc Val	tac Tyr	tcg Ser 420	cag Gln	ctc Leu	atc Ile	cgc Arg	cag Gln 425	tcc Ser	acc Thr	gcg Ala	ctc Leu	acg Thr 430	cac His	gtc Val	ttc Phe	15094				
aac Asn	cgc Arg 435	ttc Phe	ccc Pro	gag Glu	aac Asn	cag Gln 440	atc Ile	ctg Leu	gtg Val	cgc Arg	ccg Pro 445	ccc Pro	gct Ala	ccg Pro	acc Thr	15142				
att Ile 450	acc Thr	acc Thr	gtc Val	agt Ser	gaa Glu 455	aac Asn	gtt Val	ccc Pro	gcc Ala	ctc Leu 460	aca Thr	gat Asp	cac His	gga Gly	acc Thr 465	15190				
ctg Leu	ccg Pro	ctg Leu	cgc Arg	agc Ser 470	agt Ser	atc Ile	agt Ser	gga Gly	gtc Val 475	cag Gln	cgc Arg	gtg Val	acc Thr	atc Ile 480	act Thr	15238				
gac Asp	gcc Ala	cgg Arg	cga Arg 485	agg Arg	acc Thr	tgc Cys	ccc Pro	tac Tyr 490	gtg Val	cac His	aag Lys	gcc Ala	ctg Leu 495	ggc Gly	ata Ile	15286				
gtc Val	gct Ala	ccc Pro 500	aaa Lys	gtg Val	ctc Leu	tct Ser	agc Ser 505	cgc Arg	acc Thr	ttt Phe	tac Phe	caagcatgtc				15332				
cattctcatc			tcgcccgcaca		acaacaccgg			ctggggcctg			cgctcggccg		gcatgtacgg			15392				
cggcgcgaag			cggcgctcca		gcgagcacc			cgtccgcgtc			cgcggccact		accgggcccc			15452				
ctggggcgcc			cacaagcgcg		gcgtctccac			gcgcaccacc			gtcgacgacg		ccatcgacgc			15512				
cgtcgtggcc			caggccagac		gctaccgccg			gcccagtcg			acgggtggacg		ccgtcatcga			15572				
cagcgtggtg			gccgacgcgc		ggcgatacgc			tcgacgcaag			cggcgtctgc		accgccgtcg			15632				
ccgtcccacc			gccgccatgc		tggccgccag			agcggtcctg			agacgcgcgc		gccgcgtggg			15692				
ccgccgagcc			atgcgccgag		ccgcggccaa			cgccagcgcg			ggtcgcgccc		gtcgtcaggc			15752				
cgcccggcag			gccgccgccg		ccatcgccaa			cctggcccaa			ccccgcggg		gaaacgtgta			15812				
ctgggtgcga			gacgcgtcgg		gcgtgcgcgt			gccggtgcgc			acccgcccc		ctcggagtta			15872				
gaagacaaaa			agacggacga		agactgagtt			tccctgtcgt			tgccagcatg		agcaagcgca			15932				

agttcaaaga agagctgctg gaggccctcg tgcccgagat ctacggccccg gccgccgctg 15992
 ccgccgcggt ggcggacgtc aagcccgaag ttaagccccg cgcgctgaag cgggttaaaa 16052
 agcgggaaaa gaaagaggag aagcaggaag caggggttgct agacgtcgac gacggcgtgg 16112
 agttcgtgctg gtccttcgctg ccccgctcgcc ggggtgcagtg gcgggggtcgc cgcgtcaagc 16172
 tcgtcccgcg gccgggcacc gtggtgtctt tcacccccgg cctgcgttcg gccacgcgcg 16232
 gcctgaagcg cgagtacgac gaggtctatg gcgacgaaga catcctggag caggccgccc 16292
 agcagctcgg ggagtttgct tacggcaagc gcggccgcta cggggagggtg gcgctggcgc 16352
 tggaccaggg caatcccacg cccagcctca agcccgtcac gctgcagcag gtgctgccccg 16412
 tgagcgcgtc gaccgagagc aagcggggca tcaagaggga gatgggagc ctgcagccca 16472
 ccatgcaact catggtgccc aaacggcaga agctggagga cgtgctggag aacatgaaag 16532
 tggatcccag catcgagccc gaagtgaag tgcgacccat caaggaagtg ggcccgggccc 16592
 taggcgtgca gacggtggac attcagatcc ccgtgcgcgc ctcccccggt tctgccacca 16652
 ctacgacggc cgtggaggcc atggaaacgc agacggagct gcccgcgccc ttggcggcag 16712
 ccgccaccgc cgccgcggct acccgagaga tgggcatgca gaccgacccc tggtagcagt 16772
 tcgccggccc cgcccgtcgt ccacgagccc gtcggtacgc ggcgaccacc tcccggctcc 16832
 ctgactacgt cttgcatcct tccatcacgc cgacgcccgg ctaccgcgga acgaccttcc 16892
 gccccggtcg cgcgcgcacc accaccgccc gtcgtcgcac caccgcccgc cgtcgcagcc 16952
 gtcgcgcact ggctcccatc gcggttcgcc gcgtcgtccg ccgggggtcgc acgctgaccc 17012
 tgcccaccgc gcgttaccac cccagcatcg tcatttaacc tgcgctgccg ttttgagat 17072
 ggctctgacg tgccgctttc gcttccccgt tcggcactac cgaggaagat ctcgccgtag 17132
 gactggtcta gcgggcagcg gtctccgacg ccgccgccgc gcggtgcacc ggcgcagtaa 17192
 gggcggcatt ctgcccgcgc tgatccccat tatcgccgcc gccatcgggg cgatccccgg 17252
 cgtggcctcg gtggccttgc aagcagctcg caaaaattaa ataaagaagg cttgacactc 17312
 actgcctggt cctgactgtt tcatgcagac aagacatgga agacatcaat tttgcgtcgt 17372
 tggcccccg gcacggctcg cggccgttca tgggcacctg gaacgagatc ggcaccagcc 17432
 agctcaacgg gggcgctttc agttggagca gcctgtggag cggcattaaa aactttgggt 17492
 ccacgattaa gacctatggc aacaaggcgt ggaacagtag cactggtcag atgctccgcg 17552
 ataagctgaa ggaccagaac ttccagcaga aagtggtaga cggctctggcc tcgggcatca 17612
 acgggggtggt ggacctggcc aaccaggcgg tgacgaacca gatcaaccag cgtctggaga 17672
 acagccgccca gccgcccgcg gccctgcagc agcgtccgca ggtggaggag gtggaagtgg 17732
 aggagaagct gccgcccctg gagacggtgt cgccggtggg cgtgcctagc aagggggaga 17792
 agcggccgcg gcccgagctc gaggagaccc tagtgaccga gaccctggag ccgccctcgt 17852
 acgagcaggg cttgaaagag ggggcccagc ccctgcccac gaccggccc atcggaccca 17912
 tggccccgacc ggtctacggc aaggaacaca aagccgtgac gctagagctg cctccgcccg 17972

cgccccaccgt acccccgatg cccgggtccca ccctgggacac cgccgtgcct cgtccccgccg 18032

ccccgccgggt cgccgtggcc acgcccgcgc gcccagatcg cggagccaac tggcagagca 18092

ctctgaacag catcgtgggc ctgggagtga aaagcctgaa acgccgccgg tgttactatt 18152

aaagccagct aaatacccat gtgtttgtatg cgcctcctgt gtcacgccag aaaaagccag 18212

ccgagtgcg ggtcaccgcc gccgccaaga gcgccgcttt caag atg gcc acc ccc 18268
Met Ala Thr Pro
510

tcg atg atg ccg cag tgg tct tac atg cac atc gcc ggg cag gac gcc 18316
Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala Gly Gln Asp Ala
515 520 525

tcg gag tac ctg agc ccg ggc ctg gtg cag ttc gcc cgc gcc acc gac 18364
Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala Arg Ala Thr Asp
530 535 540

acg tac ttc agc ctg ggc aac aag ttt agg aac ccc acg gtg gcc ccc 18412
Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro Thr Val Ala Pro
545 550 555 560

acc cac gac gtg acg acg gac cgg tcc cag cgg ctg acg ctg cgg ttc 18460
Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu Thr Leu Arg Phe
565 570 575

gtg ccc gtc gac cgc gag gac acc gcg tac tcg tac aaa gtg cgc ttc 18508
Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr Lys Val Arg Phe
580 585 590

acg ctg gcc gtg ggc gac aac cgc gtg ctg gac atg gcc agc acg tac 18556
Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met Ala Ser Thr Tyr
595 600 605

ttt gac atc cgc ggc gtg ttg gac cgc ggt ccc agc ttc aaa ccc tac 18604
Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser Phe Lys Pro Tyr
610 615 620

tcc ggc acc gcc tac aac tcc ctg gcc ccc aag ggc gcc ccc aac ccg 18652
Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly Ala Pro Asn Pro
625 630 635 640

tca gaa tgg aag ggc tca gac aac aaa att agt gta aga ggt cag gct 18700
Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val Arg Gly Gln Ala
645 650 655

ccg ttt ttt agt aca tcc att aca aag gat ggt att caa gtg gcc act 18748
Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile Gln Val Ala Thr
660 665 670

gat act tct agc gga gct gtg tat gct aaa aag gaa tat cag cct gaa 18796
Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu Tyr Gln Pro Glu
675 680 685

cca caa gta ggg caa gaa caa tgg aac agc gaa gcc agt gat agt gat 18844
Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala Ser Asp Ser Asp
690 695 700

aaa gta gct ggt agg att cta aaa gac aca aca ccc atg ttc cct tgt 18892
Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro Met Phe Pro Cys
705 710 715 720

tac ggt tcc tac gcc aag ccc aca aat gaa cag ggg ggg caa ggc act 18940
Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly Gly Gln Gly Thr
725 730 735

aat act gta gat ctg cag ttc ttt gcc tct tca tcg gct acc tct acg 18988

Asn	Thr	Val	Asp 740	Leu	Gln	Phe	Phe	Ala 745	Ser	Ser	Ser	Ala	Thr 750	Ser	Thr	
cct Pro	aaa Lys	gcc Ala 755	gta Val	ctc Leu	tat Tyr	gcc Ala	gag Glu 760	gac Asp	gtg Val	gca Ala	ata Ile	gaa Glu 765	gca Ala	cca Pro	gac Asp	19036
acc Thr	cat His 770	ttg Leu	gtg Val	tac Tyr	aaa Lys	ccg Pro 775	gca Ala	gtt Val	aca Thr	acc Thr	acg Thr 780	acc Thr	act Thr	agt Ser	tcc Ser	19084
caa Gln 785	gac Asp	ctg Leu	cta Leu	act Thr	cag Gln 790	cag Gln	gct Ala	gct Ala	ccc Pro	aac Asn 795	cga Arg	ccc Pro	aac Asn	tac Tyr	att Ile 800	19132
ggc Gly	ttc Phe	agg Arg	gat Asp	aat Asn 805	ttt Phe	atc Ile	ggg Gly	ctc Leu	atg Met 810	tat Tyr	tac Tyr	aac Asn	tcc Ser	act Thr 815	ggc Gly	19180
aat Asn	atg Met	ggg Gly	gtt Val 820	ttg Leu	gca Ala	ggg Gly	caa Gln 825	gct Ala 825	tct Ser	cag Gln	cta Leu	aac Asn	gcg Ala 830	gtg Val	gtt Val	19228
gac Asp	ttg Leu	caa Gln 835	gac Asp	aga Arg	aac Asn	acc Thr	gag Glu 840	ctg Leu	tcc Ser	tac Tyr	cag Gln	ctc Leu 845	atg Met	ctt Leu	gat Asp	19276
gct Ala 850	ttg Leu	ggc Gly	gac Asp	cgc Arg	agt Ser	cgt Arg 855	tac Tyr	ttc Phe	tcc Ser	atg Met	tgg Trp 860	aac Asn	cag Gln	gcc Ala	gta Val	19324
gac Asp 865	agc Ser	tat Tyr	gac Asp	cct Pro	gat Asp 870	gtc Val	aga Arg	att Ile	att Ile	gaa Glu 875	aat Asn	cat His	ggg Gly	gtg Val	gag Glu 880	19372
gat Asp	gag Glu	ctg Leu	cca Pro	aac Asn 885	tac Tyr	tgt Cys	ttc Phe	ccg Pro	cta Leu 890	gga Gly	ggg Gly	tcg Ser	cta Leu	gta Val 895	act Thr	19420
gaa Glu	act Thr	tat Tyr	aca Thr 900	ggc Gly	cta Leu	tca Ser	ccc Pro	caa Gln 905	aac Asn	gga Gly	agt Ser	aac Asn	acg Thr 910	tgg Trp	aca Thr	19468
acc Thr	gac Asp	agc Ser 915	acc Thr	acc Thr	tat Tyr	gca Ala	act Thr 920	aga Arg	ggg Gly	gtg Val	gaa Glu	atc Ile 925	ggc Gly	tct Ser	ggc Gly	19516
aac Asn	atg Met 930	ttc Phe	gcc Ala	atg Met	gaa Glu	att Ile 935	aat Asn	ttg Leu	gcg Ala	gcc Ala	aat Asn 940	cta Leu	tgg Trp	agg Arg	agt Ser	19564
ttc Phe 945	ctg Leu	tac Tyr	tcc Ser	aac Asn	gtg Val 950	gcc Ala	ctg Leu	tac Tyr	ctg Leu	ccc Pro 955	gac Asp	gag Glu	tac Tyr	aag Lys	ctc Leu 960	19612
acc Thr	ccc Pro	gac Asp	aac Asn	atc Ile 965	acc Thr	ctc Leu	ccc Pro	gac Asp	aac Asn 970	aaa Lys	aac Asn	act Thr	tac Tyr	gac Asp 975	tac Tyr	19660
atg Met	aac Asn	ggc Gly	cgc Arg	gtg Val 980	gcc Ala	gcc Ala	ccc Pro	agc Ser 985	tcc Ser	ctc Leu	gac Asp	acc Thr	tac Tyr 990	gtc Val	aac Asn	19708
atc Ile	ggg Gly	gcg Ala 995	cgc Arg	tgg Trp	tcc Ser	ccc Pro	gac Asp 1000	ccc Pro	atg Met	gac Asp	aac Asn	gtc Val 1005	aac Asn	ccc Pro	ttc Phe	19756
aac Asn	cac His	cac His	cgc Arg	aac Asn	gcg Ala	gga Gly	ctg Leu	cgc Arg	tac Tyr	cgc Arg	tcc Ser	atg Met	ctg Leu	ctg Leu		19801

1010						1015						1020						
ggc	aac	ggc	cgc	tac	gta	ccc	ttc	cac	atc	caa	gtg	ccc	cag	aaa				19846
Gly	Asn	Gly	Arg	Tyr	Val	Pro	Phe	His	Ile	Gln	Val	Pro	Gln	Lys				
	1025					1030					1035							
ttc	ttc	gcc	atc	aaa	aac	ctc	ctg	ctc	ctc	ccc	ggg	tcc	tac	acc				19891
Phe	Phe	Ala	Ile	Lys	Asn	Leu	Leu	Leu	Leu	Pro	Gly	Ser	Tyr	Thr				
	1040					1045					1050							
tac	gag	tgg	aac	ttc	cgc	aag	gac	gtc	aac	atg	atc	ctc	cag	agc				19936
Tyr	Glu	Trp	Asn	Phe	Arg	Lys	Asp	Val	Asn	Met	Ile	Leu	Gln	Ser				
	1055					1060					1065							
agc	ctg	ggt	aac	gac	ctc	cgc	gtc	gac	ggg	gcc	agc	gtc	agg	ttc				19981
Ser	Leu	Gly	Asn	Asp	Leu	Arg	Val	Asp	Gly	Ala	Ser	Val	Arg	Phe				
	1070					1075					1080							
gac	agc	atc	aac	ctg	tac	gcc	aac	ttc	ttc	ccc	atg	gcc	cac	aac				20026
Asp	Ser	Ile	Asn	Leu	Tyr	Ala	Asn	Phe	Phe	Pro	Met	Ala	His	Asn				
	1085					1090					1095							
acc	gcc	tcc	acg	ctc	gag	gcc	atg	ctg	cgc	aac	gac	acc	aac	gac				20071
Thr	Ala	Ser	Thr	Leu	Glu	Ala	Met	Leu	Arg	Asn	Asp	Thr	Asn	Asp				
	1100					1105					1110							
cag	tcg	ttc	aac	gac	tac	ctc	tgc	gct	gcc	aac	atg	ctc	tac	ccc				20116
Gln	Ser	Phe	Asn	Asp	Tyr	Leu	Cys	Ala	Ala	Asn	Met	Leu	Tyr	Pro				
	1115					1120					1125							
atc	ccc	gcc	aac	gcc	acc	agc	gtg	ccc	atc	tcc	att	ccc	tcg	cgg				20161
Ile	Pro	Ala	Asn	Ala	Thr	Ser	Val	Pro	Ile	Ser	Ile	Pro	Ser	Arg				
	1130					1135					1140							
aac	tgg	gcc	gcc	ttc	cgg	ggc	tgg	agc	ttc	acc	cgg	ctc	aag	acc				20206
Asn	Trp	Ala	Ala	Phe	Arg	Gly	Trp	Ser	Phe	Thr	Arg	Leu	Lys	Thr				
	1145					1150					1155							
aag	gag	acc	ccc	tct	ctg	ggc	tcc	ggc	ttc	gat	ccc	tac	ttc	acc				20251
Lys	Glu	Thr	Pro	Ser	Leu	Gly	Ser	Gly	Phe	Asp	Pro	Tyr	Phe	Thr				
	1160					1165					1170							
tac	tcg	ggc	tcc	atc	ccc	tac	ctg	gac	ggc	acc	ttc	tac	ctc	aac				20296
Tyr	Ser	Gly	Ser	Ile	Pro	Tyr	Leu	Asp	Gly	Thr	Phe	Tyr	Leu	Asn				
	1175					1180					1185							
cac	act	ttc	aag	aag	gtc	tcc	atc	atg	ttc	gac	tcc	tcc	gtc	agc				20341
His	Thr	Phe	Lys	Lys	Val	Ser	Ile	Met	Phe	Asp	Ser	Ser	Val	Ser				
	1190					1195					1200							
tgg	ccc	ggc	aac	gac	cgc	ctg	ctg	acc	ccc	aac	gag	ttc	gag	atc				20386
Trp	Pro	Gly	Asn	Asp	Arg	Leu	Leu	Thr	Pro	Asn	Glu	Phe	Glu	Ile				
	1205					1210					1215							
aag	cgc	acc	gtg	gac	ggg	gaa	ggg	tac	aac	gtg	gcc	cag	tgc	aac				20431
Lys	Arg	Thr	Val	Asp	Gly	Glu	Gly	Tyr	Asn	Val	Ala	Gln	Cys	Asn				
	1220					1225					1230							
atg	acc	aag	gac	tgg	ttc	ctc	atc	cag	atg	ctc	agc	cac	tac	aac				20476
Met	Thr	Lys	Asp	Trp	Phe	Leu	Ile	Gln	Met	Leu	Ser	His	Tyr	Asn				
	1235					1240					1245							
atc	ggc	tac	cag	ggc	ttc	tac	gtg	ccc	gag	ggc	tac	aag	gac	agg				20521
Ile	Gly	Tyr	Gln	Gly	Phe	Tyr	Val	Pro	Glu	Gly	Tyr	Lys	Asp	Arg				
	1250					1255					1260							
atg	tac	tct	ttc	ttc	cgc	aac	ttc	caa	ccc	atg	agc	cgc	cag	gtg				20566
Met	Tyr	Ser	Phe	Phe	Arg	Asn	Phe	Gln	Pro	Met	Ser	Arg	Gln	Val				
	1265					1270					1275							

gtc gac acc acc acc tac acc gac tac aaa aac gtc acc ctc ccc Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu Pro	20611
1280 1285 1290	
ttc cag cac aac aac tcg ggg ttc gtg gga tac atg ggc ccc acc Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr	20656
1295 1300 1305	
atg cgc gag ggg cag gcc tac ccc gcc aac tac ccc tac ccc ctg Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu	20701
1310 1315 1320	
atc ggc aag acc gcc gtg ccc agc ctc acg cag aaa aag ttc ctc Ile Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu	20746
1325 1330 1335	
tgc gac cgc acc atg tgg cgc atc ccc ttc tcc agt aac ttc atg Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met	20791
1340 1345 1350	
tcc atg ggg gcg ctc acc gac ctg ggg cag aac atg ctg tac gcc Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala	20836
1355 1360 1365	
aac tcc gcc cac gcc ctc gac atg acc ttc gag gtg gac ccc atg Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met	20881
1370 1375 1380	
gat gag ccc acg ctt ctc tat gtt ctg ttc gaa gtg ttc gac gtc Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val Phe Asp Val	20926
1385 1390 1395	
gtg cgc atc cac cag ccg cac cgc ggc gtc atc gag gcc gtc tac Val Arg Ile His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr	20971
1400 1405 1410	
ctg cgc acg ccg ttc tcg gcc ggt aac gcc acc acc taa ggagggggcc Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr	21020
1415 1420 1425	
gccgacggat gggctccagc gagccggagc tggctgccat cgcgcgcgac ctgggctgcg	21080
ggccctactt cctgggcacc ttgacaaac gcttcccggg cttcgtggcg ccgcacaagc	21140
tggcctgcgc catcgtcaac accgccggac gcgagaccgg cggcgtccac tggctggccc	21200
tggcctggaa cccccgcagc cgaacctgct acctcttcga ccccttcggc ttctcggacg	21260
acaggctcag gcagatctac cagttcgagt acgaaggcct gctccggcgc agcgccctcg	21320
cctccacccc cgaccactgc gtcaccctcg tcaagtccac ccagaccgtc caggggcccc	21380
gctcggccgc ctgcggcctc ttctgctgca tgttcctgca cgccttcgtg cgctggcccc	21440
cctcccccat ggacggcaac cccaccatgg acctccttac gggcgttccc aacagcatgc	21500
ttcagagtcc ccaggctcag cccaccctcc accgcaacca ggaggaactc tacgccttcc	21560
tggctcggca ctccccctac tttcgccgcc accgcgagcg catagaaaag gccaccgcgt	21620
ttgacaaaat gaacgactag attttctgtg aaaaacactc aataaagcct ttattggttc	21680
accacacgtg cacgcatgca gactttttat ttaaaagggc tccgcctcct cgtcgccgtg	21740
gctggtgggg agggagacgt tgcgatactg caggcggggag ctccatctga actcgggaat	21800
cagcagcttg ggcagggggc cctcgacgtt ctgcgtccac agcttgcgca ccagctgcag	21860

ggcgcccagc aggtcgggcg cggagatctt gaagtcgcag ttggggccct ggttgccgcg 21920
 ggagttgcgg tacaccgggt tggcgactg gaacaccagc acgctgggggt gctcgatgct 21980
 ggccagcgcc gtcttgtcgg tcacctcgtc gccgcgcagg gactccgcgt tgctcagcgc 22040
 gaaggcggtc agcttgcaca gctgccgacc cagcacgggc accccgctcg gctggttcag 22100
 gcagtcgcag cgcatagcca tcagcagccg cttctgcccg tgctgcatct tcggatagtc 22160
 ggctcgcagt aaggcctcca tctgccggaa ggccgtctgc gccttgctgc cctccgagaa 22220
 gaacagcccc caggacttgc cggagaacac gttgttgccg cagctcacgt cttccacgca 22280
 gcagcgcgcg tcgtcgttct tcagctgcac cacgctgcgg cccagcgggt tctgcaccac 22340
 cttggtcttg ccgggatgtt cttcagggc ccgctggccg ttctcgttg tcacgtccat 22400
 ctccaccacc tgctccttct ggatcatctc cagcccgttg tagcagcgca gcacgccctc 22460
 ctgctcgggt caccctgca gccagacggc gcagccggtc ggctccagct gttgaggttt 22520
 caccgccgcg taggtctcca cgtacgccg caggaagcgg cccatcatct ccacaaaggt 22580
 cttctgaccg gtgaagggtc gctgcagccc gcgatgctcc tcgttgagcc acgtctgaca 22640
 gatcttgccg tacaccttgc cctgctcggg cagaaacttg aaagcggcct tctcctcggg 22700
 ctccacgttg tacttctcca tcagcgccga catcagctcc atgcccttct cccaggccga 22760
 caccagcggc tccgcgcggg ggttcaccac cgccatgcct cgggaagtgc cggggcgctc 22820
 atcttctctc tcctcctcgt cttcttcttg aggcggcgggt ggcggcagtt gtctcacgaa 22880
 tctcttgccg ttggccttct ggacgatctc cacgccgggg tgggtgaacc cgtgggccac 22940
 caccattcgt tcctcttctt cttcgtgtc gggcacgact tcgggagagg gaggcggcgg 23000
 aggaaccgggt gcggccactg cggccatcgc ggcgttcttg cgcgcttct tggggggcag 23060
 aggcggcgtc tcgcgtccg ggctggtctc ttgcaggtag ggcgtgatgg tgtgggaggt 23120
 ggggcgctct ggctgacggc cggccatgct gatgcttgac tcctaggcga aaagatggag 23180
 gaggatctta gacagccgca gcccgctctc gaaaccttaa ccacccccgc ctctgaggtc 23240
 ggcgccggcg agctagacat gcaacgggag gaggaggagg acgtgcgagt ggagcaagac 23300
 ccgggctacg tgacgccg ccaggacggc gaggagccgc aggcaccggc gccaacgctc 23360
 agcgaagccg actacctggg aggggaggac gacgtgctgc tgaagcacct ggcgcggcag 23420
 agcaccatcg tgcaggaggc cctcaaggag cgcgaggagg tcccgtgac ggtggaggag 23480
 ctacagccgg cctacgaagc caacctctt tcgccgcggg tgcccccaa gaagcaggcc 23540
 aacggcacct gcgagcccaa ccccgccctc aacttctacc ccgtctttgc ggtgcccag 23600
 gcgctggcca cctatcacat cttcttcaag aaccagcgca tccccctctc gtgccgcgcc 23660
 aaccgaccc gcgccgacc cctcctgcat ctccgagccg gcgccgcat acctgagatc 23720
 gcctccctgg aggaagtccc caagatcttc gaaggtctcg gcaaggacga gaagcgcgcg 23780
 gcaaacgctc tggaaaagaa cgagagcgag ggtcagaacg tgctggtcga gctggaaggc 23840
 gacaacgcgc gtctggccgt gctcaaacgc accatcgaag tctcccactt cgcctacccc 23900
 gcgctcaacc ttcccccaa ggtcatgcgc tcggatcatg atcagctgct catcaagcgc 23960

gccgagcccc	tcgagaacga	ctccgaggtg	gattccgagg	acggaaaacc	cgtggtctcg	24020
gacgaggagc	tcgcgcgctg	gctgggcacg	caggaccccc	ccgagttgca	agagcggcgc	24080
aagatgatga	tggcggccgt	gctggtcacc	gccgagctcg	agtgcctgca	gcgcttcttc	24140
gccgaccccc	agaccctgcg	caaggtcgag	gagtccttgc	actacgcctt	ccgccacggc	24200
tacgtgcgcc	aggcctgcaa	gatctccaac	gtggagctta	gcaacctggg	ctcctacatg	24260
ggcatcctgc	acgagaaccg	cctcgggcag	aacgtcctcc	actgcaccct	gaccggggag	24320
gcccgcgcg	actacgtccg	cgactgcac	tacctctttc	tcacctcac	ctggcagacc	24380
gcatggggg	tctggcagca	gtgtctggag	gagcgcaacc	ttcgcgagct	cgacaagcta	24440
ctgagccgcg	agcgcgcga	gctctggacg	gctttcagcg	agcgcaccgc	cgcttgccgt	24500
ctggccgacc	tcattctccc	cgagcgactc	aggcaaacc	tccagaacgg	cctgcccga	24560
tttgtcagcc	agagcatgct	gcaaaacttt	cgctccttca	tcctggagcg	atccggcatc	24620
ttgcccgcga	tgagctgcgc	cctgccctcc	gatttcgtcc	ccctctatta	tcgcgagtgc	24680
ccccgcgcg	tctggagcca	ctgctacctg	ctgcgtctgg	ccaactacct	cgcccaccac	24740
tccgacctca	tgaagactc	cagcggcgag	gggctgctgg	agtgccactg	ccgctgcaac	24800
ctctgcaccc	cccaccgctc	gctggtctgc	aacaccgagc	tgctcagcga	gacgcaagtg	24860
atcggtacct	ttgagatcca	gggaccagag	gggcccggag	gtgcttccaa	cctcaagctc	24920
agcccggcgc	tctggacttc	cgctacctg	cgcaaattta	tccccgagga	ctatcacgcc	24980
caccagatcc	aattctacga	agaccaatcg	cgacccccca	aagccccctt	cacggcctgt	25040
gtcatcaccc	agagccagat	tctggcccaa	ttgcaagcca	tccagcaggc	ccgccaagag	25100
ttcctcctga	aaaagggtca	cggggtctat	ctggaccccc	agaccggcga	ggaactcaac	25160
accccgtcac	cctccgccgc	cgcttcgtgc	cgcccgcaga	accatgccgc	ccaaagggaa	25220
caagcaggcc	atcgcccagc	ggcgggcca	gaagcagcaa	gagctccagg	agcagtggga	25280
cgaggagtcc	tgggacagcc	aggcggagga	agtctcagac	gaggaggagg	acatggagag	25340
ctgggacagc	ctagacgagg	aggaggaggc	cgaggagcta	gaggacgagc	ctctcgagga	25400
ggaagagccc	agcagcgccg	cggcaccatc	ggcttccaaa	gaagcggctc	ggagccggcc	25460
ggccccgaag	cagcagaagc	agcaacagcc	gccaccgtcg	cccccgacgc	caccaccagg	25520
ctcactcaaa	gccagccgta	ggtgggacgc	ggtgtccatc	gcgggatcgc	ccaaagcccc	25580
agtcggtaag	ccaccggggc	ggtcgcggcg	ggggtactgt	tcctggcgcc	cccacaagag	25640
caagatcgtc	gcctgcctcc	agcactgccg	gggcaacatc	tccttcgcgc	ggcgctactt	25700
gctcttccac	gacgggggtg	cggtgccgcg	caacgtcctc	tactattacc	gtcatctcta	25760
cagcccctac	gagacagaag	gcccggcctc	cgcgtaagac	cagccgccag	acggtctcct	25820
ccgccatcgc	gaccgcagc	gactcggccg	ccacgcagga	gctcagaaaa	cgcatctttc	25880
ccaccctgta	tgctatcttc	cagcagagcc	gcggccagca	gctggaactg	aaagtaaaaa	25940
accgctccct	gcgttcgctc	acccgcagct	gtctgtacca	caggagggaa	gaccaactgc	26000

agcgcacgct cgaggacgcc gaggcactgt tcaataaata ctgctcgggtg tctcttaagg 26060
 actgaaagcc cgcgcttttt cagaggctca ttacgtcatc atcatcatga gcaaggacat 26120
 tcccacgcct tacatgtgga gctaccagcc gcagatggga ctggcgggccg gcgcctccca 26180
 ggattactcc agtcgcatga actggctgag tgccggcccc cacatgatcg ggcggggtcaa 26240
 tgggattcgt gccacccgca atcagatact gctggaacag gccgccctca cctccacccc 26300
 gcgacgtcag ctgaacccgc ccgcttggcc cgccgcccag gtgtaccagg aaaaccccg 26360
 cccgaccaca gtcctcctgc cacgcgacgc ggaggccgaa gtccagatga ctaactccgg 26420
 ggcgcaatta gcgggcggcg cccgccacgt cgtcgctccc gggtagagag gtcggccccgc 26480
 accctacccc tccggccccta taaagaggct gatcattcga ggccgaggta tccagctcaa 26540
 cgacgaggtg gtgagctcct cgaccggtct tcggcccgac ggagtcttcc agcttggagg 26600
 cgccggccgc tcttccttca ccactcgcca ggccctacctg acgctccaga gctcttcctc 26660
 ccagcctcgc tccggcgcca tcggcaccct ccagttcgtg gaggagtctg tgccctcggt 26720
 ctacttcaac ccgttctccg gctctcccg cgctacccg gacagcttca tccccaaacta 26780
 cgacgcggtg agcgaatccg tggacggcta cgattgatga ccgatgggtgc ggccgtaact 26840
 gcgcggcggc aacatctgca tcaactgccat cgctcctcgggt gcttcgcccc ggaggcctgt 26900
 gagttcatct acttccagct cgccccggac cagcttcagg gcccttcgca cggcggttaag 26960
 ctcgtgatag aggaagagct cgagagtagc tgccctgcgt gttttacctc gcgccccatc 27020
 ctagtcgaga gggaaacgcgg taggaccacc ctcaccctct actgcatctg tgactccccg 27080
 gaattacatg aagatctgtg ttgccttcta tgtgccgaac aataaccctt cttgtaacta 27140
 cctacatcca caataaacca gaatttggaa actcctttcg tttgtttgca g atg aaa 27197
 Met Lys

cgc gcc cgc ctc gac gac gac ttc aac ccc gtc tac ccc tat gac 27242
 Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr Asp
 1430 1435 1440
 act ccc aac gct ccc tct gtt ccc ttc atc act cct ccc ttc gtc 27287
 Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val
 1445 1450 1455
 tcc tcg gac ggc ttg caa gaa aaa cca ccc gga atg ctc agt ctc 27332
 Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu
 1460 1465 1470
 aac tac caa gat cct att acc acc caa aac ggg gca tta act cta 27377
 Asn Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu
 1475 1480 1485
 aag ctt ggc agc gga ctg aac ata aac caa gat ggg gaa ctt acc 27422
 Lys Leu Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr
 1490 1495 1500
 tca gac gcc agc gtt ctc gtc act ccc ccc att aca aaa gcc aac 27467
 Ser Asp Ala Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn
 1505 1510 1515
 aac aca ata ggc cta gcc ttc aat gca cct ctt acc ttg caa agc 27512
 Asn Thr Ile Gly Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser
 1520 1525 1530

gat	act	tta	aat	ctt	gct	tgt	aac	gcc	cca	ctt	acc	gtg	caa	gac	27557
Asp	Thr	Leu	Asn	Leu	Ala	Cys	Asn	Ala	Pro	Leu	Thr	Val	Gln	Asp	
		1535					1540					1545			
aat	agg	ttg	gga	ata	aca	tac	aac	tct	ccc	ctc	acc	ttg	caa	aac	27602
Asn	Arg	Leu	Gly	Ile	Thr	Tyr	Asn	Ser	Pro	Leu	Thr	Leu	Gln	Asn	
		1550					1555					1560			
agc	gaa	ctt	gcc	cta	gcg	gtc	acc	ccg	cct	ctt	gac	act	gcc	aat	27647
Ser	Glu	Leu	Ala	Leu	Ala	Val	Thr	Pro	Pro	Leu	Asp	Thr	Ala	Asn	
		1565					1570					1575			
aac	aca	ctt	gcg	ctt	aaa	acc	gcc	cgg	cct	ata	att	aca	aac	tct	27692
Asn	Thr	Leu	Ala	Leu	Lys	Thr	Ala	Arg	Pro	Ile	Ile	Thr	Asn	Ser	
		1580					1585					1590			
aat	aac	gag	ctt	aca	ctc	tcc	gct	gat	gct	ccc	cta	aac	acc	agc	27737
Asn	Asn	Glu	Leu	Thr	Leu	Ser	Ala	Asp	Ala	Pro	Leu	Asn	Thr	Ser	
		1595					1600					1605			
acg	ggc	acc	ctc	cgc	cta	caa	agc	gca	gca	cca	ctg	ggg	cta	gtt	27782
Thr	Gly	Thr	Leu	Arg	Leu	Gln	Ser	Ala	Ala	Pro	Leu	Gly	Leu	Val	
		1610					1615					1620			
gac	caa	acc	ctg	cga	gtg	ctt	ttt	tct	aac	cca	ctc	tac	ttg	caa	27827
Asp	Gln	Thr	Leu	Arg	Val	Leu	Phe	Ser	Asn	Pro	Leu	Tyr	Leu	Gln	
		1625					1630					1635			
aac	aac	ttt	ctc	tca	cta	gcc	att	gaa	cgc	cca	ttg	gct	tta	act	27872
Asn	Asn	Phe	Leu	Ser	Leu	Ala	Ile	Glu	Arg	Pro	Leu	Ala	Leu	Thr	
		1640					1645					1650			
acc	act	ggc	tct	atg	gct	atg	cag	att	tcc	caa	cca	tta	aaa	gtg	27917
Thr	Thr	Gly	Ser	Met	Ala	Met	Gln	Ile	Ser	Gln	Pro	Leu	Lys	Val	
		1655					1660					1665			
gaa	gac	gga	agc	tta	agc	ttg	agc	att	gaa	agc	cct	cta	aat	cta	27962
Glu	Asp	Gly	Ser	Leu	Ser	Leu	Ser	Ile	Glu	Ser	Pro	Leu	Asn	Leu	
		1670					1675					1680			
aaa	aac	gga	aat	ctt	act	tta	gga	acc	caa	agt	ccc	cta	act	gtc	28007
Lys	Asn	Gly	Asn	Leu	Thr	Leu	Gly	Thr	Gln	Ser	Pro	Leu	Thr	Val	
		1685					1690					1695			
act	ggc	aac	aac	ctc	agc	ctt	aca	aca	aca	gcc	cca	tta	acg	gtt	28052
Thr	Gly	Asn	Asn	Leu	Ser	Leu	Thr	Thr	Thr	Ala	Pro	Leu	Thr	Val	
		1700					1705					1710			
cag	aac	aac	gct	cta	gcc	ctc	tca	gtg	tta	ctg	ccg	ctt	aga	cta	28097
Gln	Asn	Asn	Ala	Leu	Ala	Leu	Ser	Val	Leu	Leu	Pro	Leu	Arg	Leu	
		1715					1720					1725			
ttt	aat	aac	acc	tca	ctg	gga	gtg	gca	ttc	aac	cca	ccc	att	tct	28142
Phe	Asn	Asn	Thr	Ser	Leu	Gly	Val	Ala	Phe	Asn	Pro	Pro	Ile	Ser	
		1730					1735					1740			
tca	gca	aac	aac	ggg	ctg	tct	ctt	gac	att	gga	aat	ggc	ctt	aca	28187
Ser	Ala	Asn	Asn	Gly	Leu	Ser	Leu	Asp	Ile	Gly	Asn	Gly	Leu	Thr	
		1745					1750					1755			
ctg	caa	tac	aac	agg	ctc	gta	gtg	aac	att	ggc	ggc	ggg	cta	cag	28232
Leu	Gln	Tyr	Asn	Arg	Leu	Val	Val	Asn	Ile	Gly	Gly	Gly	Leu	Gln	
		1760					1765					1770			
ttt	aac	aac	ggc	gct	att	acc	gct	tcc	ata	aat	gca	gct	ctg	ccg	28277
Phe	Asn	Asn	Gly	Ala	Ile	Thr	Ala	Ser	Ile	Asn	Ala	Ala	Leu	Pro	
		1775					1780					1785			

ttg cag tat tcc aat aac cag ctt tct ctt aat att gga ggc ggg	28322
Leu Gln Tyr Ser Asn Asn Gln Leu Ser Leu Asn Ile Gly Gly Gly	
1790 1795 1800	
ctg cga tac aac ggc act tac aaa aat tta gcc gtc aaa acc gac	28367
Leu Arg Tyr Asn Gly Thr Tyr Lys Asn Leu Ala Val Lys Thr Asp	
1805 1810 1815	
tct ttt agg ggt ctt gaa att gac agt aat cag ttc ctg gtg cca	28412
Ser Phe Arg Gly Leu Glu Ile Asp Ser Asn Gln Phe Leu Val Pro	
1820 1825 1830	
aga ctg ggt tct ggt cta aag ttt gat caa tat ggg tac att agc	28457
Arg Leu Gly Ser Gly Leu Lys Phe Asp Gln Tyr Gly Tyr Ile Ser	
1835 1840 1845	
gtc ata cct cca act gtt acg cca aca aca ctt tgg act aca gca	28502
Val Ile Pro Pro Thr Val Thr Pro Thr Thr Leu Trp Thr Thr Ala	
1850 1855 1860	
gac cct tct ccc aac gct act ttt tac gac agc tta gat gct aag	28547
Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser Leu Asp Ala Lys	
1865 1870 1875	
gta tgg ctg gcc tta gta aaa tgc aac ggc atg gtt aat gga acc	28592
Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val Asn Gly Thr	
1880 1885 1890	
ata gcc ata aag gct tta aaa ggt act ctg ctc caa cct acg gct	28637
Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro Thr Ala	
1895 1900 1905	
agt ttt att tct ttt gtt atg tat ttt tac agc aat ggc acc aga	28682
Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr Arg	
1910 1915 1920	
aga act aac tac ccc acg ttt gaa aat gaa ggc ata cta gct agt	28727
Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser	
1925 1930 1935	
agt gct aca tgg ggt tat cgt caa gga aac tcg gca aac acc aac	28772
Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn	
1940 1945 1950	
gtc acc agt gcc gtt gaa ttt atg cct agc tcc aca aga tat cct	28817
Val Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro	
1955 1960 1965	
gtt aac aag ggt act gag gtt cag aac atg gaa ctc acc tac act	28862
Val Asn Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr	
1970 1975 1980	
ttc ttg cag gga gac ccc act atg gcc ata tca ttt caa gct att	28907
Phe Leu Gln Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile	
1985 1990 1995	
tat aac cat gct ttg gaa ggt tac tct tta aaa ttt acc tgg cga	28952
Tyr Asn His Ala Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg	
2000 2005 2010	
gtt cgc aac agg gaa cgc ttt gat atc ccc tgc tgt tct ttt tct	28997
Val Arg Asn Arg Glu Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser	
2015 2020 2025	
tac ata acg gaa gaa taa acactgtttt tcttttcaat gtttttattc	29045
Tyr Ile Thr Glu Glu	
2030	

tgctttttta cacagttcga accgtcagac tccctccccc ctccacttc acccggtaca 29105

cctcccgctc cccctggatc gctgcgtaca actgcagttt ggtgttcaga cacgggttct 29165
taggtgacag tatccacacg gcctctttgc cggccaggcg ctggtccgta atgctcacia 29225
atccctccga cacgtcctcc agacacacgg tggaatccaa ggcgcccgtc tacaaaacaa 29285
acacagtcac gctctccacg ggttctctcc tcggtcgtac tgcgccagcg tgaacgggcg 29345
atggtgctcc atcagggtc gcagcaaccg ctgtcggcgc ggctcaccca ggctccggcg 29405
aaaagcgccc cgtctgggag tgctattcaa aaaacgcacc gcctttatca acagtctcct 29465
cgtgcggcg ggcagcagc gcacctggat ctctgtcagg tctttacaat aggtacagcc 29525
catcaccacc atgttgttta aaatcccaaa gctaaacacg ctccacccaa atgacatgaa 29585
ttccagcacc gccgcggcgt ggccatcata caatatgcgg aggtaaatca ggtgccgccc 29645
cctaatacaa acgctcccca tatacatcac ctcttaggc agttgataat taaccacctc 29705
ccggtaccag ggaaacctca cgtttactaa agcccaaac accaacattt taaaccagtt 29765
agccagcacc accctcccg ccttactg cagcgacccc ggctgtttac aatgacagt 29825
aatcacccac ctctcatacc ccctaattgac ctggcggtgc tccacatcta tagtagcaca 29885
gcacacgcac accctcatgt aatgcttcat cacaaatctt tccaagggg ttagtatcat 29945
gtcccagggt acgggccact cctgcagcac ggtgaaagggt acgcaggcg gaaacagtcct 30005
cacctcggac acataatgca tattcagatg ttcacactct aaaaccccg ggcttcctc 30065
caacgcagcc actggcaagt tctcagaggg tgggtgaagg cggtggtgct gatagggact 30125
caatctgtgt cgacaccgtc tgtcgcgttg catcgtagac caacgcttg cgacaccgct 30185
cgtacttcgc ccaaagaaaa cgggtgcgac gccaacacac ttccgcgtac cgtgggttcc 30245
gcactcgagc tcgctcagtt ctcaacgcac aatgcagcca ttctgtaat ccacacaaca 30305
gtcgtcggc ttccaaagag atgtgcacct cgtatcttat aacgtcccga tatatatcca 30365
agcaggcagt cagggccact tgcaaccagt gcacgcaggc ggactgatcg cgacacactg 30425
gaggtggagg gagagacgga agaggcatgt tactccagac ggtcgaaaag cggatcaaag 30485
tgcagatcgc gaagatggca gcgatcccc cgcgtacgct ggtgatagat cacagccagg 30545
tcaaacataa tgcggttttc caaatgacct attaccgcct ccaccagagc cgccacgcgc 30605
acttcagaa acaccagcac ggctacggca ttctctcaa aatcttcaa cattaagctg 30665
catgattgaa tcaccccaa ataattctcc tcttccatt ctcgcaaat ttgagtaaaa 30725
acctctcgca gattagctcc gtggcggtca aaaagggtcac ttagagcgcc ctccaccgcc 30785
atgcgcaagc acaccctcat gattgaaaaa tgccagtctc ctgaaccacc tgcagttgat 30845
ttaaaagacc tatattagga tcaattccac tctccgcag ctccacgcgt agcattagct 30905
gcaaaaagtc atttaaatct tcgcaacta gcgcggtaag ctgcgcccg ggaattaggt 30965
ctgaagcagt caccacacac ataatttcca gtgaaggagt cagtctaagc agcaaaaagc 31025
cgcatgagca gtgttgaaaa ggaggggtca cgcaatgtaa catatgcagc caaaaatctc 31085
caagggtgtc gtgcataaac tccaccactg aaaagtccaa atcatgtaaa tatgccatca 31145

ccgcctcagg aaccaccacg gacacaaaaa cgggccgtag caaatacatg gtgtcctgca 31205
 aagcaaaaac acatttatac catagaggcg cgaattactt ggggaaaaat cactcgctcc 31265
 aaaactaaac aggccaccgt ctgaccgcgc cagccataaa aaaagcggtt cgaatgatta 31325
 aaaagaataa tagacacctc ccaccaggtta ctcggctgca actcgtgcgc ccctatcaaa 31385
 accccgcgga cgttcatgtc ggccatagaa aaaatgcggc ccaaatatcc caccggaatc 31445
 tccacggcca gctgcagtga tagcaaaaga acgccatgag gagcaatcac aaaattttca 31505
 ggcgataaaa gcacataaag gttagaatag ccctgctgca caggtaataa agcccgcgag 31565
 ctcagcaaat gcacataaac cgcttcagcc atcccgtctt accgcgaaca aaaggctcac 31625
 agtacacagt tactcaacct acacgccaca cagtatttat aactcctca atcgccacgt 31685
 caccgcggcc gaacaaactc caaaagtcca aaaagtccaa aacgcccgcg taaaagcccg 31745
 ccaaaacagc acttcctcat ttactctccc acagtacgtc acttccgccc cgcccgcgcg 31805
 cctcgccccg cctcaccct cgcgtccac ccgcgcccc acgtcagact cccaccccgc 31865
 cccgcgcccc cgctatccgc accccaccct cactccaccc ctaaccccgc ctcctcatta 31925
 tcatattggc accgttccca aataaggtat attatgatga tg 31967

<210> 13
 <211> 508
 <212> PRT
 <213> Simian adenovirus

<400> 13

Met Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly
 1 5 10 15

Pro Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala
 20 25 30

Thr Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu
 35 40 45

Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr
 50 55 60

Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu
 65 70 75 80

Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn
 85 90 95

Asn Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp
 100 105 110

Glu Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn
 115 120 125

Met Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg
 Page 124

130	135	140
Leu Met Val Glu Lys Val 145	Asn Lys Glu Thr Asn 150	Ala Pro Arg Tyr Glu 155
Trp Phe Glu Phe Thr 165	Leu Pro Glu Gly Asn 170	Tyr Ser Glu Thr Met Thr 175
Ile Asp Leu Met Asn Asn Ala Ile Val 180	Asp Asn Tyr Leu Glu Val Gly 185	
Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr 195	200	205
Arg Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro 210	215	220
Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro 225	230	235
Gly Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly 245	250	255
Ile Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu 260	265	270
Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr 275	280	285
Glu Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly 290	295	300
Asp Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu 305	310	315
Lys Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr 325	330	335
Asn Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro 340	345	350
Glu Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr 355	360	365
Cys Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp 370	375	380
Pro Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val 385	390	395
Gly Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln 405	410	415

Ala Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val
 420 425 430

Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro
 435 440 445

Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly
 450 455 460

Thr Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile
 465 470 475 480

Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly
 485 490 495

Ile Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe
 500 505

<210> 14
 <211> 917
 <212> PRT
 <213> Simian adenovirus

<400> 14

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala
 1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
 20 25 30

Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
 35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
 50 55 60

Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
 65 70 75 80

Lys Val Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
 85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
 100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
 115 120 125

Ala Pro Asn Pro Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val
 130 135 140

Arg Gly Gln Ala Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile
145 150 155 160

Gln Val Ala Thr Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu
165 170 175

Tyr Gln Pro Glu Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala
180 185 190

Ser Asp Ser Asp Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro
195 200 205

Met Phe Pro Cys Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly
210 215 220

Gly Gln Gly Thr Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser
225 230 235 240

Ala Thr Ser Thr Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile
245 250 255

Glu Ala Pro Asp Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr
260 265 270

Thr Thr Ser Ser Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg
275 280 285

Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr
290 295 300

Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
305 310 315 320

Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln
325 330 335

Leu Met Leu Asp Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp
340 345 350

Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
355 360 365

His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly
370 375 380

Ser Leu Val Thr Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser
385 390 395 400

Asn Thr Trp Thr Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu
405 410 415

Ile Gly Ser Gly Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn
Page 127

420	425	430
Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp		
435	440	445
Glu Tyr Lys Leu Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn		
450	455	460
Thr Tyr Asp Tyr Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp		
465	470	475
Thr Tyr Val Asn Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn		
485	490	495
Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser		
500	505	510
Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro		
515	520	525
Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr		
530	535	540
Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser		
545	550	555
Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Val Arg Phe Asp		
565	570	575
Ser Ile Asn Leu Tyr Ala Asn Phe Phe Pro Met Ala His Asn Thr Ala		
580	585	590
Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe		
595	600	605
Asn Asp Tyr Leu Cys Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn		
610	615	620
Ala Thr Ser Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe		
625	630	635
Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu		
645	650	655
Gly Ser Gly Phe Asp Pro Tyr Phe Thr Tyr Ser Gly Ser Ile Pro Tyr		
660	665	670
Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile		
675	680	685
Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr		
690	695	700

Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
705 710 715 720

Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu
725 730 735

Ser His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
740 745 750

Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
755 760 765

Gln Val Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu
770 775 780

Pro Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr
785 790 795 800

Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
805 810 815

Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu Cys Asp
820 825 830

Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
835 840 845

Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
850 855 860

Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu
865 870 875 880

Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Ile His Gln Pro
885 890 895

His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
900 905 910

Gly Asn Ala Thr Thr
915

<210> 15
<211> 607
<212> PRT
<213> Simian adenovirus

<400> 15

Met Lys Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr
1 5 10 15

Asp Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val
 20 25 30
 Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu Asn
 35 40 45
 Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu Lys Leu
 50 55 60
 Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr Ser Asp Ala
 65 70 75 80
 Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn Asn Thr Ile Gly
 85 90 95
 Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser Asp Thr Leu Asn Leu
 100 105 110
 Ala Cys Asn Ala Pro Leu Thr Val Gln Asp Asn Arg Leu Gly Ile Thr
 115 120 125
 Tyr Asn Ser Pro Leu Thr Leu Gln Asn Ser Glu Leu Ala Leu Ala Val
 130 135 140
 Thr Pro Pro Leu Asp Thr Ala Asn Asn Thr Leu Ala Leu Lys Thr Ala
 145 150 155 160
 Arg Pro Ile Ile Thr Asn Ser Asn Asn Glu Leu Thr Leu Ser Ala Asp
 165 170 175
 Ala Pro Leu Asn Thr Ser Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala
 180 185 190
 Pro Leu Gly Leu Val Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro
 195 200 205
 Leu Tyr Leu Gln Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu
 210 215 220
 Ala Leu Thr Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu
 225 230 235 240
 Lys Val Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn
 245 250 255
 Leu Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val
 260 265 270
 Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val Gln
 275 280 285
 Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu Phe Asn

290	295	300
Asn Thr Ser Leu Gly Val 305	Ala Phe Asn Pro 310	Pro Ile Ser Ser Ala Asn 315 320
Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr Leu Gln Tyr Asn 325		330 335
Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln Phe Asn Asn Gly Ala 340		345 350
Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro Leu Gln Tyr Ser Asn Asn 355		360 365
Gln Leu Ser Leu Asn Ile Gly Gly Gly Leu Arg Tyr Asn Gly Thr Tyr 370		375 380
Lys Asn Leu Ala Val Lys Thr Asp Ser Phe Arg Gly Leu Glu Ile Asp 385		390 395 400
Ser Asn Gln Phe Leu Val Pro Arg Leu Gly Ser Gly Leu Lys Phe Asp 405		410 415
Gln Tyr Gly Tyr Ile Ser Val Ile Pro Pro Thr Val Thr Pro Thr Thr 420		425 430
Leu Trp Thr Thr Ala Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser 435		440 445
Leu Asp Ala Lys Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val 450		455 460
Asn Gly Thr Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro 465		470 475 480
Thr Ala Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr 485		490 495
Arg Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser 500		505 510
Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn Val 515		520 525
Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro Val Asn 530		535 540
Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr Phe Leu Gln 545		550 555 560
Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile Tyr Asn His Ala 565		570 575

Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg Val Arg Asn Arg Glu
 580 585 590

Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser Tyr Ile Thr Glu Glu
 595 600 605

<210> 16
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic oligomer

<400> 16
 gcgacgggcc gacgctgccc ggct

24

<210> 17
 <211> 4
 <212> PRT
 <213> Artificial

<220>
 <223> artificial

<400> 17

Arg Arg Ala Ser
 1

<210> 18
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic oligomer

<400> 18
 gcggcgcgcc gacgctgccc ggct

24